IN THE UNITED STATES DISTRICT COURT

FOR THE EASTERN DISTRICT OF VIRGINIA

Alexandria Division

DIALECT, LLC,
    Plaintiff,

                                                            REDACTED

        v.
                                                            Civil No. 1:23cv581 (DJN)

AMAZON.COM INC., *et al.*,
    Defendants.

MEMORANDUM OPINION

This patent infringement case centers on Plaintiff Dialect, LLC's allegations that Amazon

Alexa, a virtual assistant product developed by Defendants Amazon.com, Inc. and Amazon Web

Services, Inc. (together, "Amazon"), infringes six of Dialect's patents.[1] The parties have

completed discovery, and Amazon now moves for summary judgment on all claims. Amazon

presents two theories that it claims warrant summary judgment. First, Amazon argues that Alexa

does not infringe on Dialect's patents as a matter of law. Second, Amazon renews its contention,

first raised at the motion to dismiss stage, that five of the six Asserted Patents do not satisfy 35

U.S.C. § 101's requirement of patentable subject matter. The parties have fully briefed

Amazon's Motion. (ECF Nos. 234, 255, 278, 331, 357.) The Court now stands ready to rule.[2]

---

[1]   U.S. Pat. Nos. 7,693,720 (the "'720 Patent"); 8,015,006 (the "'006 Patent"); 8,140,327
(the "'327 Patent"); 8,195,468 (the "'468 Patent"); 9,263,039 (the "'039 Patent"); and 9,495,957
(the "'957 Patent") (together, the "Asserted Patents"). In general, this Memorandum Opinion
refers to patents using their last three digits, as it does for the Asserted Patents.

[2]   On May 17, 2024, Amazon moved for oral argument on its motion for summary
judgment. (ECF No. 236.) Because the parties have thoroughly presented their arguments in the
briefs, the Court does not find that oral argument would assist in resolving Amazon's Motion.
Accordingly, the Court will deny Amazon's motion for oral argument. Loc. Civ. R. 7(J); Fed. R.
Civ. P. 78(b).

After describing its reasoning, the Court concludes that it must grant Amazon's motion for summary judgment in part and deny it in part. Alexa does not infringe the '327 Patent and parts of the '006 Patent as a matter of law, and none of the Asserted Patents claim abstract ideas in violation of 35 U.S.C. § 101. Those disputes that remain must be resolved at trial and by a jury.

I.   BACKGROUND

The facts giving rise to this case have been recounted in prior opinions. *Dialect, LLC v. Amazon.com, Inc.* (the "*Alice* Opinion"), __ F. Supp. 3d __, 2023 WL 7381551 (E.D. Va. Nov. 8, 2023) (ECF No. 58); *Dialect, LLC v. Amazon.com, Inc.* (the "*Markman* Opinion"), 2024 WL 1859806 (E.D. Va. Apr. 29, 2024) (ECF No. 212). The Court now briefly surveys the relevant facts, which neither party disputes for purposes of summary judgment.

Dialect owns a suite of software patents originally developed by the defunct start-up VoiceBox Technologies. Founded in 2001, VoiceBox developed marketable speech recognition and natural language processing devices until it was acquired by Nuance Communications in 2018.[3] Over its 17-year lifespan, VoiceBox acquired a sizable portfolio of patents for itself. Later, several of those patents — the Asserted Patents among them — were assigned to Dialect.

Seeking to monetize its VoiceBox patents, Dialect sued Amazon on May 1, 2023, and the case was initially assigned to Senior District Judge T.S. Ellis, III. Dialect's Complaint (and later, its Amended Complaint) pleaded seven counts of infringement predicated on seven patents. Amazon moved to dismiss six of those counts on the ground that the patent claims asserted were directed to patent-ineligible subject matter.

---

[3]   Taylor Soper, *Nuance Buys Voicebox Technologies, Scooping up Speech-Recognition and Natural-Language Pioneer*, GeekWire (May 18, 2018, 10:54 AM), https://www.geekwire.com/2018/nuance-communications-buys-voicebox-technologies-scooping-another-seattle-area-company/ [https://perma.cc/69DS-82LQ].

Judge Ellis granted Amazon's motion to dismiss in part and deferred it in part. (ECF No. 59.) First, Judge Ellis found that the claim asserted in Count V of Dialect's Amended Complaint — Claim 1 of U.S. Patent No. 9,031,845 — was directed to the abstract idea of "using context to execute a spoken request." *Alice Op.*, 2023 WL 7381551, at *4 (applying 35 U.S.C. § 101 and *Alice Corp. v. CLS Bank Int'l*, 573 U.S. 208 (2014)). Judge Ellis reasoned that the process disclosed by Claim 1 of the '845 Patent, which he described as "understanding language using context, determining whether an on-or off-board processor is to handle that language, and then using that processor to execute the language" was "no less abstract than 'collection of information, comprehension of its meaning, and indication of the results,'" an idea repeatedly recognized as abstract in Federal Circuit case law. *Id.* (alterations omitted). Next, Judge Ellis found that Dialect failed to plead any facts suggesting that Claim 1 contained an inventive concept that would save it from ineligibility. *Id.* at *5–6. Accordingly, he found Claim 1 invalid as a matter of law. *Id.* at *6. Judge Ellis did not analyze any of the five other patents challenged by Amazon's motion to dismiss. *Id.* at *1. In his view, the issue was best deferred, because "claim construction and discovery may helpfully inform the [eligibility] analysis." *Id.* at *6. He therefore opted to reevaluate the subject-matter eligibility of Amazon's five remaining challenged patents following "further proceedings." *Id.*

Following Judge Ellis's *Alice* Opinion, discovery began. In January 2024, the matter was transferred to the undersigned. (ECF No. 137.) The parties then briefed their respective constructions of disputed claim terms in the Asserted Patents. At the same time, in what was effectively an early motion for summary judgment, Amazon sought to have several claims of the Asserted Patents declared indefinite and therefore invalid as a matter of law. On April 29, 2024,

3

the Court issued a *Markman* Order, which construed the parties' disputed terms and rejected

Amazon's indefiniteness challenges. (ECF No. 213.)

At the parties' claim construction hearing, the Court instructed that summary judgment in

this case would be bifurcated, with claim construction–dependent grounds for summary

judgment (such as noninfringement, indefiniteness and subject-matter eligibility) to be briefed

first and claim construction–independent grounds (such as novelty, obviousness and enablement)

to be briefed second. (ECF No. 205 ¶ 8(a)–(b).) Since then, the parties completed discovery,

and on May 16, 2024, Amazon filed the instant Motion, which asserts two grounds for summary

judgment — noninfringement and subject-matter invalidity.[4]

II.   THE ASSERTED PATENTS

To properly contextualize the analysis that follows, the Court begins by describing the six

patents at issue in Amazon's Motion.

A.   The '006 Patent

The '006 Patent continues U.S. Patent No. 7,398,209.[5] '006 Pat. [63]. The '006 Patent,

titled "Systems and Methods for Processing Natural Language Speech Utterances With Context-

Specific Domain Agents," discloses "a fully integrated environment allowing users to submit

natural language speech questions and commands." *Id.* at col. 1, ll. 22–24. Dialect asserts

Claims 1, 2, 3, 5, 10 and 11 of the '006 Patent. (Amazon's Br. at 1; Dialect's Opp. at 1.) Of

---

[4]   The Court's *Markman* Opinion held that Amazon's indefiniteness arguments failed to
rebut the presumption of validity that attached to the Asserted Patents. The Court allowed
Amazon to renew its indefiniteness arguments on summary judgment if it could meet its burden
of proof. 2024 WL 1859806, at *22. Amazon has not taken up that invitation in this Motion, so
the Court does not address indefiniteness in its Memorandum Opinion.

[5]   A "continuation" patent discloses additional claims on an invention patented by a
previous application. *See* 35 U.S.C. § 120 (describing "[a]n application for patent for an
invention . . . in an application previously filed in the United States, . . . which is filed by an
inventor or inventors named in the previously filed application").

those six claims, three are independent — Claims 1, 5, and 10.[6] Claim 1 adequately represents

all three independent claims asserted, so the Court describes only Claim 1 in detail.

Claim 1 of the '006 Patent is directed to a method for processing natural speech.[7] In

plain language, the claimed method consists of the following steps:

(a) Capturing audio that contains a spoken request by using a "speech unit";

---

[6]   A "dependent" claim incorporates a prior claim; an independent claim does not. *See* 35
U.S.C. § 112(d) ("A claim in dependent form shall be construed to incorporate by reference all
the limitations of the claim to which it refers.").

[7]   In full, Claim 1 reads as follows:

"A method for processing natural language speech utterances with context-specific domain agents,
comprising:"

(a) "receiving, at a speech unit coupled to a processing device, a natural language speech
utterance that contains a request;"

(b) "recognizing, at a speech recognition engine coupled to the processing device, one or more
words or phrases contained in the utterance using information in one or more dictionary and
phrase tables, wherein recognizing the one or more words or phrases contained in the
utterance includes:"

(1) "dynamically updating the information in the one or more dictionary and phrase tables
based on a dynamic set of prior probabilities or fuzzy possibilities;"

(2) "determining an identity associated with a user that spoke the utterance based on voice
characteristics associated with the utterance; and"

(3) "associating the one or more recognized words or phrases and a pronunciation associated
with the one or more recognized words or phrases with the determined identity and the
request contained in the utterance"

(A) "in response to the one or more recognized words or phrases satisfying a
predetermined confidence level;"

(c) "parsing, at a parser coupled to the processing device, the one or more recognized words or
phrases to determine a meaning associated with the utterance and a context associated with
the request contained in the utterance,"

(1) "wherein the one or more recognized words or phrases are further associated with the
determined context in response to the one or more recognized words or phrases satisfying
the predetermined confidence level;"

(d) "formulating, at the parser, the request contained in the utterance in accordance with a
grammar used by a domain agent associated with the determined context;"

(e) "processing the formulated request with the domain agent associated with the determined
context to generate a response to the utterance; and"

(f) "presenting the generated response to the utterance via the speech unit."

'006 Pat. col. 25, l. 49–col. 26, l. 21.

450

FIG. 4B

(b)   Identifying words contained in that audio using a database of dictionary entries by:

(1)   Updating the dictionary entries using certain mathematical techniques;

(2)   Analyzing the audio of the spoken request to determine the request's speaker; and

(3) Associating the identified words and the speaker's pronunciation of those words, on the one hand, with the speaker's identity and the meaning of the spoken request, on the other hand,

(A) Though only when the identified words first "satisfy[] a predetermined confidence level; and then

(c) Parsing the identified words to determine the meaning of the audio and the context of the request encoded in that audio,

(1) Such that, when the identified words satisfy the "confidence level" of step (b)(3)(A), those words are associated with the context in question;

(d) Using the context to pick out the right agent to handle the request and then "translating" the request into a form usable by the selected agent;

(e) Using the selected agent to generate a response to the spoken request; and finally

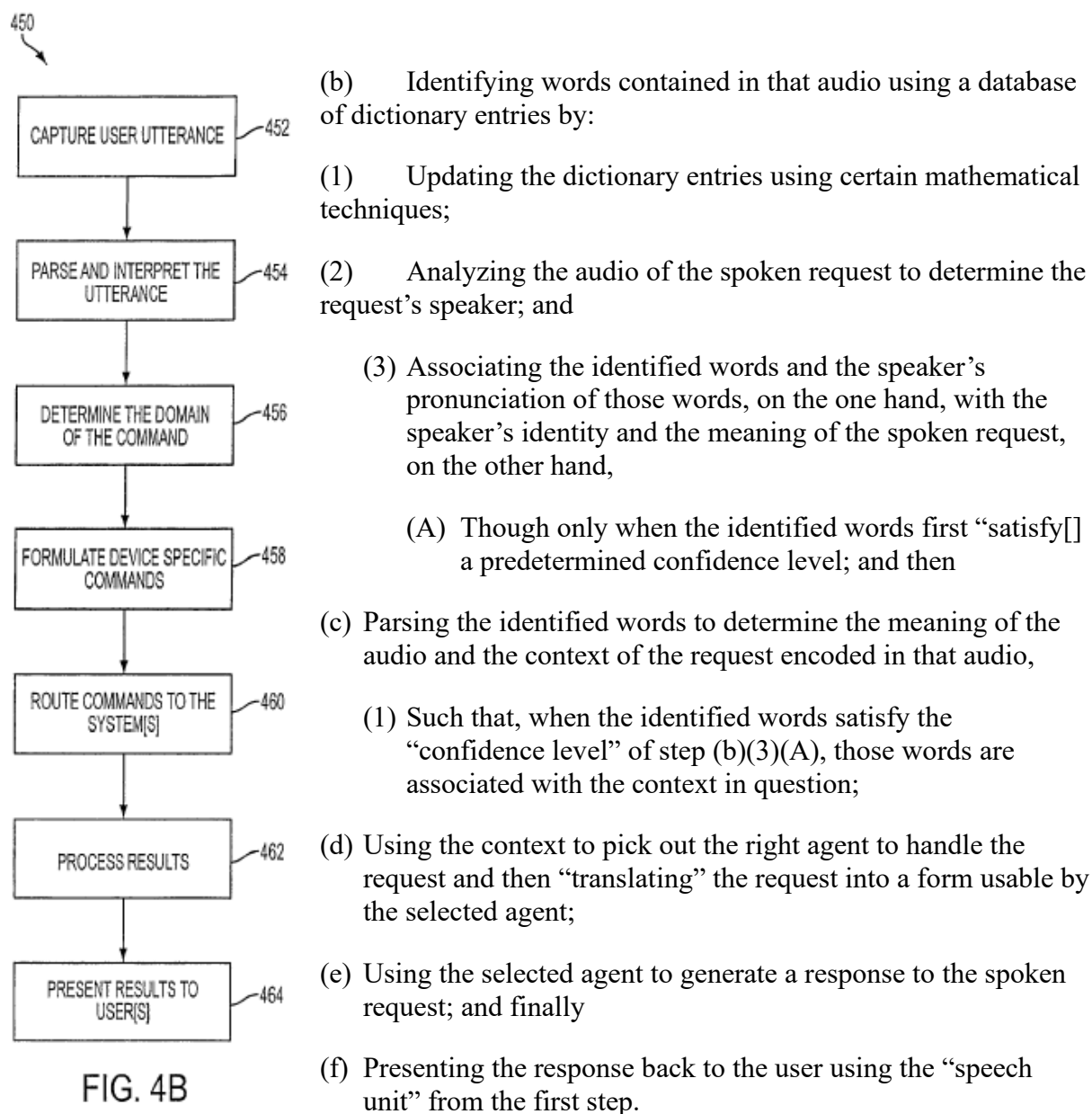(f) Presenting the response back to the user using the "speech unit" from the first step.

Claims 5 and 10 contain the same basic limitations as Claim 1. Claim 1 provides greater detail on the "identifying" process employed used in step (b), Claim 5 provides detail on the "translating" process used in step (d) and Claim 10 provides detail on the "presenting" process set out in step (f). All three method claims have the same basic structure, which is reproduced in the left-hand margin of this page. '006 Pat. fig.4B.

B.   The '720 Patent

The '720 Patent is titled "Mobile Systems and Methods for Responding to Natural

Language Speech Utterance" and, much like the '006 Patent, discloses "a fully integrated

environment allowing mobile users to ask natural language questions or give natural language

commands." *Id.* at col. 1, ll. 15–17. Dialect asserts Claims 1, 4, 14, 19, 31 and 32 of the '720

Patent. (ECF No. 235 ("Amazon's Br.") at 1; ECF No. 256 ("Dialect's Opp.") at 1.)   Claim 1,[8]

---

[8]   In full, Claim 1 reads as follows:

"A mobile system responsive to a user generated natural language speech utterance, comprising:"

(a) "a speech unit connected to a computer device on a vehicle, wherein the speech unit receives a natural language speech utterance from a user and converts the received natural language speech utterance into an electronic signal; and"

(b) "a natural language speech processing system connected to the computer device on the vehicle, wherein the natural language speech processing system receives, processes, and responds to the electronic signal using data received from a plurality of domain agents, wherein the natural language speech processing system includes:"

(1) "a speech recognition engine that recognizes at least one of words or phrases from the electronic signal using at least the data received from the plurality of domain agents, wherein the data used by the speech recognition engine includes a plurality of dictionary and phrase entries that are dynamically updated based on at least a history of a current dialog and one or more prior dialogs associated with the user,"

(2) "a parser that interprets the recognized words or phrases, wherein the parser uses at least the data received from the plurality of domain agents to interpret the recognized words or phrases, wherein the parser interprets the recognized words or phrases by:"

(A) "determining a context for the natural language speech utterance;"

(B) "selecting at least one of the plurality of domain agents based on the determined context; and"

(C) "transforming the recognized words or phrases into at least one of a question or a command, wherein the at least one question or command is formulated in a grammar that the selected domain agent uses to process the formulated question or command; and"

(c) "an agent architecture that communicatively couples services of each of [(1)] an agent manager, [(2)] a system agent, the plurality of domain agents, and [(3)] an agent library that includes one or more utilities that can be used by the system agent and the plurality of domain agents, wherein the selected domain agent uses the communicatively coupled services to create a response to the formulated question or command and format the response for presentation to the user."

'720 Pat. col. 32, l. 30–col. 33, l. 6.

the only truly independent claim asserted,[9] is directed to a computer architecture comprising three elements. Stated in simple terms, the claim may be summarized as follows:

(a) The first claim element is a generic computer program on a vehicle that transforms spoken language into an electronic signal.

(b) The second element is a system of generic programs that processes and responds to the electronic signal. This system includes the following sub-elements:

   (1) A generic speech recognition engine that recognizes discrete words or phrases from the signal using data received from a plurality of dynamically updated entries in an electronic database, or dictionary. Those dynamically updated dictionary entries must be based on both the ongoing dialog and one or more prior dialogs with the same user.

   (2) A generic parser that assigns meaning to the words and phrases generated by the speech recognition engine by first

      (A) determining a context for the words and phrases, then

      (B) selecting one or more sub-processors based on that context, and finally

      (C) using the words and phrases to generate a question or command formulated in machine language that can be understood by the selected sub-processors.

(c) The third element is an architecture that binds together the foregoing processors together with (1) an agent manager, (2) a system agent, and (3) an electronic library containing tools usable by the system to create a response to the spoken language and present it to the user.

   A representative embodiment of such a system is included in the '720 Patent's specification, as pictured below:

---

[9]   Claim 1 is the only truly independent asserted claim, because the other independent claim asserted, Claim 31, simply restates Claim 1's "system" as a "method." *See* '720 Pat. col. 35, ll. 23–61 (Claim 31).
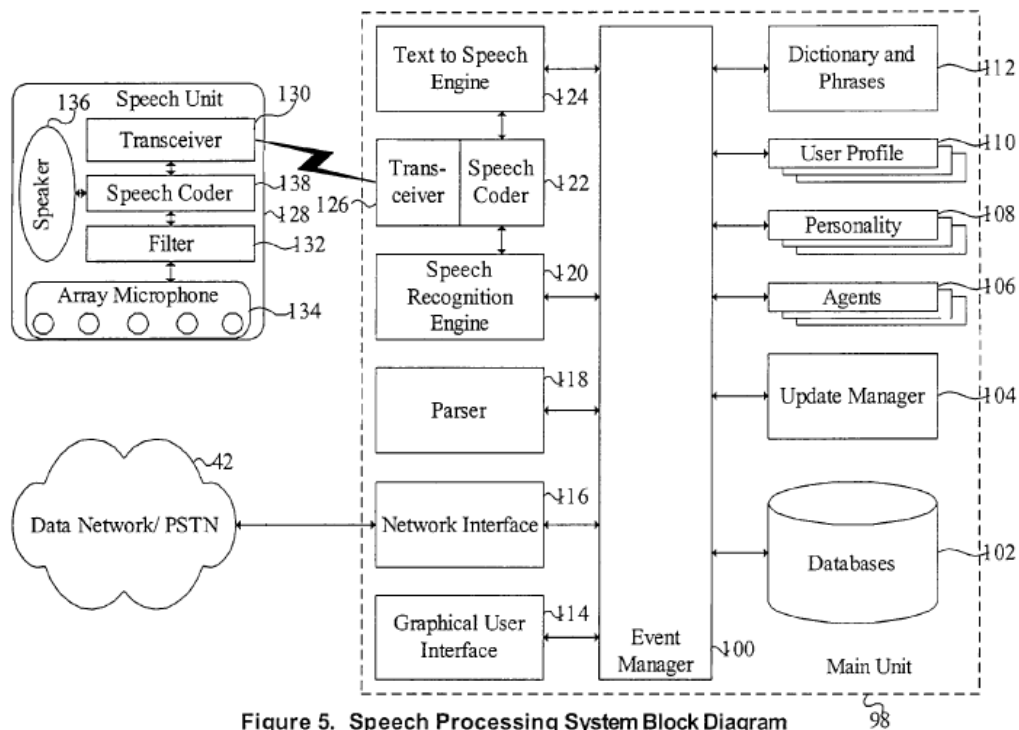
Figure 5. Speech Processing System Block Diagram

'720 Pat. fig.5. The elements that compose the claimed computer architecture are defined

mostly — but not entirely — in terms of their function.

C.   The '957 Patent

The '957 Patent constitutes the third continuation of the '468 Patent, which itself results

from a division[10] of U.S. Patent No. 7,949,529 — a VoiceBox patent filed on August 29, 2005,

and issued on May 24, 2011. The '957 Patent, like the '529 Patent and its progeny, is titled

"Mobile Systems and Methods of Supporting Natural Language Human-Machine Interactions"

and discloses a series of systems that "enables mobile users to submit natural language speech

and/or non-speech questions or commands" to a mobile computer. *Id.* at col. 1, ll. 28–29.

Dialect asserts Claims 1, 3, 4, 5, 7 and 8 of the '957 Patent. (Amazon's Br. at 1; Dialect's Opp.

---

[10]   "If two or more independent and distinct inventions are claimed in one application," an
inventor may make the second invention "the subject of a divisional application." 35 U.S.C.
§ 121.

at 1.) Claim 1 of that patent,[11] the only truly independent claim asserted, is directed to a system

for processing natural speech.[12] Claim 1 of the '957 Patent is composed of elements defined

only in functional terms. In plain language, Claim 1 is directed to any system whose processors

can perform the following steps:

> (a) Create a "context stack," i.e., a database of context entries generated based on prior
>   spoken language;
>
> (b) Receive a spoken message that is associated with a command or request;
>
> (c) Break that message down into words;
>
> (d) Identify a context entry corresponding to the message by (1) comparing the entries to the
>   words of the message, (2) generating "rank scores" based on each comparison, and
>   (3) using those rank scores to select the appropriate context entry in the context stack;
>   and

---

[11]   In full, Claim 1 reads as follows:

"A system for processing a natural language utterance, the system including one or more
processors executing one or more computer program modules which, when executed, cause the
one or more processors to:"

> (a) "generate a context stack comprising context information that corresponds to a plurality of
>   prior utterances, wherein the context stack includes a plurality of context entries;"
>
> (b) "receive the natural language utterance, wherein the natural language utterance is associated
>   with a command or is associated with a request;"
>
> (c) "determine one or more words of the natural language utterance by performing speech
>   recognition on the natural language utterance;"
>
> (d) "identify, from among the plurality of context entries, one or more context entries that
>   correspond to the one or more words, wherein the context information includes the one or
>   more context entries, wherein identifying the one or more context entries comprises:"
>
>> (1) "comparing the plurality of context entries to the one or more words;"
>>
>> (2) "generating, based on the comparison, one or more rank scores for individual context
>>   entries of the plurality of context entries; and"
>>
>> (3) "identifying, based on the one or more rank scores, the one or more context entries from
>>   among the plurality of context entries; and"
>
> (e) "determine, based on the determined one or more words and the context information, the
>   command or the request associated with the natural language utterance."

'957 Pat. col. 39, ll. 38–67.

[12]   As with the '720 Patent, Claim 1 of the '957 Patent proves to be the only *truly*
independent claim because the other independent claim asserted, Claim 7, does no more than
restate Claim 1's "system" as a "method." *See* '957 Pat. col. 40, l. 42–col. 41, l. 3 (Claim 7).

(e) Use the selected context entry to determine which pre-programmed command or request corresponds to the message.

D.   The '468 Patent

The '468 Patent, like the '957 Patent, traces its lineage back to VoiceBox's application for the '529 Patent. '468 Pat. col. 1, ll. 8–13. Like the '529 Patent and the '957 Patent, the '468 Patent is titled "Mobile Systems and Methods of Supporting Natural Language Human-Machine Interactions" and purportedly "enables mobile users to submit natural language speech and/or non-speech questions or commands" to a mobile computer. *Id.* at col. 1, ll. 28–29. Dialect asserts Claims 19, 20, 26, 27, 28, 29, 30 and 32 of the '468 Patent. (Amazon's Br. at 1; Dialect's Opp. at 1.) Claim 19,[13] the only independent claim asserted, is directed to a method for

---

[13]   In full, Claim 19 reads as follows:

"A method for processing multi-modal natural language inputs, comprising:"

(a) "receiving a multi-modal natural language input at a conversational voice user interface, the multi-modal input including a natural language utterance and a non-speech input provided by a user, wherein a transcription module coupled to the conversational voice user interface transcribes the non-speech input to create a non-speech-based transcription;"

(b) "identifying the user that provided the multi-modal input;"

(c) "creating a speech-based transcription of the natural language utterance using a speech recognition engine and a semantic knowledge-based model, wherein the semantic knowledge-based model includes

   (1) "a personalized cognitive model derived from one or more prior interactions between the identified user and the conversational voice user interface,"

   (2) "a general cognitive model derived from one or more prior interactions between a plurality of users and the conversational voice user interface, and"

   (3) "an environmental model derived from an environment of the identified user and the conversational voice user interface;"

(d) "merging the speech-based transcription and the non-speech-based transcription to create a merged transcription;"

(e) "identifying one or more entries in a context stack matching information contained in the merged transcription;"

(f) "determining a most likely context for the multi-modal input based on the identified entries;"

(g) "identifying a domain agent associated with the most likely context for the multi-modal input;"

(h) "communicating a request to the identified domain agent; and"

processing "multi-modal" inputs. The claim specifies that a "multi-modal natural language input" consists of a combination of "a natural language utterance" — i.e., spoken words — and "a non-speech input."[14] *Id.* at col. 40, ll. 40–42. In simplified terms, the claimed method comprises the following steps:

    (a) Receiving a multi-modal input (speech and non-speech input) and directly transcribing the non-speech component of the input;

    (b) Identifying the user that provided the input;

    (c) Transcribing the spoken component of the input using a speech recognition engine and a "semantic knowledge-based model" that includes three elements:

        (1) a "personalized" model based on prior interactions with the identified user,

        (2) a "general" model based on prior interactions with other users, and

        (3) an "environmental" model;

    (d) Combining the non-speech transcription and the speech transcription into a single transcription;

    (e) Identifying entries in a context stack that match the resulting transcription;

    (f) Using those entries to determine the most likely context for the multi-modal input;

    (g) Using that context to determine which domain agent (i.e., which sub-program) should respond to that input;

    (h) Referring the multi-modal input to the appropriate sub-program; and

    (*i*) Delivering the result back to the user.

A helpful, though simplified, diagram of the claimed method is included in the '468 Patent's specification and is reproduced below:

        (*i*) "generating a response to the user from content provided by the identified domain agent as a result of processing the request."

'468 Pat. col. 40, l. 37–col. 41, l. 5.

[14]   At oral argument on Amazon's motion to dismiss, counsel for Dialect suggested that pressing a button, typing on a keyboard, gesturing on a touch screen or moving a mouse would constitute common forms of "non-speech input." (ECF No. 80 (Hr'g Tr.) at 12, 27.)
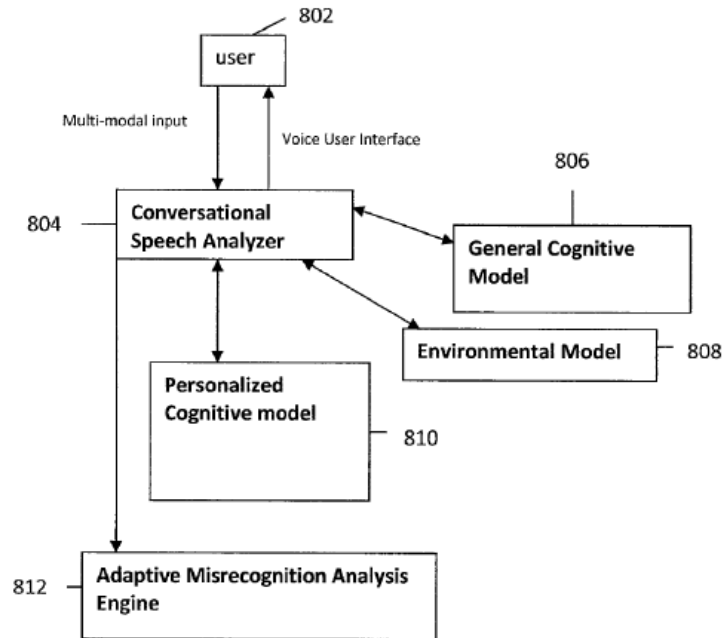
FIGURE 8

'468 Pat. fig.8.

E.   The '039 Patent

The '039 Patent constitutes the fourth continuation of another VoiceBox patent, U.S.

Patent No. 7,640,160. '039 Pat. col. 1, ll. 8–20. The '039 Patent is titled "Systems and Methods

for Responding to Natural Language Speech Utterance" and discloses, in its own words, "a fully

integrated environment that allows users to submit natural language questions and commands"

through, as relevant here, "a combination of a speech interface and a non-speech interface." *Id.*

at col. 1, ll. 25–31. Dialect asserts Claims 13, 14, 15, 16, 17 and 18 of the '039 Patent.

(Amazon's Br. at 1; Dialect's Opp. at 1.) The only independent claim of these six, Claim 13,[15] is

---

[15]   That claim reads as follows:

"A method of processing speech and non-speech communications, comprising:"

(a) "receiving the speech and non-speech communications;"

13

directed to "a method of processing speech and non-speech communications" (i.e., a method of processing what the '468 Patent calls "multi-modal inputs"). *Id.* at col. 30, ll. 39–40. The claimed method may be summarized as consisting of the following steps:

(a) Receiving a multi-modal input;

(b) Transcribing the speech and non-speech components of that input separately;

(c) Merging the speech-and non-speech transcriptions into a single query;

(d) Looking for various "text combinations" — e.g., a particular character or word or word group (*see id.* at col. 13, ll. 61–65) — within that query;

(e) Comparing those combinations to entries in a "context description grammar";

(f) Invoking sub-programs (or "domain agents") associated with that grammar;

(g) Generating a "relevance score" representing the result of the comparison;

(h) Choosing the right sub-program to execute the query based on the relevance score;

(*i*) Retrieving the content of that sub-program; and

(j) Returning the results to the user, with the results arranged based on the relevance score.

(b) "transcribing the speech and non-speech communications to create a speech-based textual message and a non-speech-based textual message;"

(c) "merging the speech-based textual message and the non-speech-based textual message to generate a query;"

(d) "searching the query for text combinations;"

(e) "comparing the text combinations to entries in a context description grammar;"

(f) "accessing a plurality of domain agents that are associated with the context description grammar;"

(g) "generating a relevance score based on results from comparing the text combinations to entries in the context description grammar;"

(h) "selecting one or more domain agents based on results from the relevance score;"

(*i*) "obtaining content that is gathered by the selected domain agents; and"

(j) "generating a response from the content, wherein the content is arranged in a selected order based on results from the relevance score."

'039 Pat. col. 30, ll. 39–61.

Put even more simply, Claim 13 appears to describe a method for interpreting a multi-modal command using comparison, scoring and iteration.

      F.   **The '327 Patent**

The '327 Patent, the last of the six Asserted Patents, results from a division of U.S. Patent No. 7,809,570, which itself resulted from a division of the '209 Patent. '327 Pat. [62]. The '327 Patent bears the title "System and Method for Filtering and Eliminating Noise from Natural Language Utterances to Improve Speech Recognition and Parsing." '327 Pat. [12], [45]. Much like other inventions at issue in this case, the invention disclosed by the '327 Patent constitutes "a fully integrated environment allowing users to submit natural language speech questions and commands." *Id.* at col. 1, ll. 27–29.

Dialect asserts that Amazon Alexa infringes Claims 1, 5, 6, 14, 18 and 19 of the '327 Patent. (Amazon's Br. at 1; Dialect's Opp. at 1.) Claim 1,[16] the only truly independent claim, is

---

[16]   For reference, Claim 1 of the '039 Patent reads as follows:

"A method for filtering and eliminating noise from natural language utterances, comprising:"

   (a) "receiving a natural language utterance at a microphone array that adds one or more nulls to a beam pattern steered to point in a direction associated with a user speaking the natural language utterance, wherein the one or more nulls notch out point or limited area noise sources from an input speech signal corresponding to the natural language utterance;"

   (b) "comparing environmental noise to the input speech signal corresponding to the natural language utterance to set one or more parameters associated with an adaptive filter coupled to the microphone array;"

   (c) "passing the input speech signal corresponding to the natural language utterance to the adaptive filter, wherein the adaptive filter uses band shaping and notch filtering to remove narrow-band noise from the input speech signal corresponding to the natural language utterance according to the one or more parameters;"

   (d) "suppressing cross-talk and environmentally caused echoes in the input speech signal corresponding to the natural language utterance using adaptive echo cancellation in the adaptive filter;"

   (e) "sending the input speech signal passed through the adaptive filter to a speech coder that uses adaptive lossy audio compression to remove momentary gaps from the input speech signal and variable rate sampling to compress and digitize the input speech signal, wherein the speech coder optimizes the adaptive lossy audio compression and the variable rate sampling to only preserve components in the input speech signal that will be input to a speech recognition engine; and"

directed to "[a] method for filtering an eliminating noise from natural language utterances."[17] *Id.*

at col. 25, ll. 59–60. Unlike the other Asserted Patents, the '327 Patent does not face an

eligibility challenge brought by Amazon's Motion. Accordingly, for present purposes, it suffices

to note that the claimed method recites a specific series of technological steps to achieve its goal

of "filter[ing] and eliminat[ing] noise from natural language utterances to improve accuracy

associated with speech recognition." '327 Pat. [57]. Amazon argues that its Alexa technology

does not feature several of the specific steps recited in the claims and, accordingly, does not

infringe the '327 Patent as a matter of law.

III.   LEGAL STANDARDS

Federal Rule of Civil Procedure 56 requires the Court to grant summary judgment to

Amazon "if [it] shows that there is no genuine dispute of material fact and the movant is entitled

to judgment as a matter of law." Fed. R. Civ. P. 56(a). The party with the burden of proof at trial

must point to evidence that, if believed by a reasonable factfinder, would warrant ruling in that

party's favor. *Celotex Corp. v. Catrett*, 477 U.S. 317, 322–23 (1986). In other words, Amazon

need not "produce evidence showing the absence of a genuine issue of material fact" with

respect to all issues; instead, where Dialect bears the burden of proof, Amazon need only point to

"an absence of evidence to support [Dialect's] case." *Id.* at 325.

(f) "transmitting the digitized input speech signal from a buffer in the speech coder to the speech
recognition engine, wherein the speech coder transmits the digitized input speech signal to the
speech recognition engine at a rate that depends on available bandwidth between the speech
coder and the speech recognition engine."

'327 Pat. col. 25, l. 59–col. 26, l. 30.

[17]   As before, although Claim 14 of the '327 Patent is independent in form, it restates Claim
1's "method" as a "system" and thus lacks substantive independence. *See* '327 Pat. col. 27,
ll. 14–54 (Claim 14).

When determining whether a genuine issue has been created, all "justifiable inferences are to be drawn in the [nonmovant's] favor." *Anderson v. Liberty Lobby, Inc.*, 477 U.S. 242, 255 (1986). In patent cases, "the law of the regional circuit" — here, the Fourth Circuit — determines the summary judgment standard. *ADASA Inc. v. Avery Dennison Corp.*, 55 F.4th 900, 907 (Fed. Cir. 2022). Thus, binding precedent counsels that in this posture, the Court "may not credit [Amazon's] evidence, weigh the evidence, or resolve factual disputes in [Amazon's] favor." *Hensley ex rel. North Carolina v. Price*, 876 F.3d 573, 579 (4th Cir. 2017). However, Dialect must do more than engage in "mere speculation or the building of one inference upon another." *Othentec Ltd. v. Phelan*, 526 F.3d 135, 140 (4th Cir. 2008). Instead, it must successfully present evidence such that "a jury applying [the relevant] evidentiary standard could reasonably find" in Dialect's favor. *Anderson*, 477 U.S. at 255.

The Patent Act provides that any person who "makes, uses, offers to sell, or sells any patented invention . . . infringes the patent." 35 U.S.C. § 271(a). "The patentee has the burden of proving infringement by a preponderance of the evidence." *Eli Lilly & Co. v. Hospira, Inc.*, 933 F.3d 1320, 1328 (Fed. Cir. 2019). Determining whether a patent claim has been infringed requires a two-step analysis. *Tessera, Inc. v. Int'l Trade Comm'n*, 646 F.3d 1357, 1364 (Fed. Cir. 2011). First, the claim must be construed. Then, the proper construction must be applied to the accused products. That second step constitutes "a question of fact." *Id.* For that reason, the ultimate question of infringement "is amenable to summary judgment when no reasonable factfinder could find that the accused product contains every claim limitation or its equivalent." *Akzo Nobel Coatings, Inc. v. Dow Chem. Co.*, 811 F.3d 1334, 1339 (Fed. Cir. 2016). If even a single claim limitation "is totally missing from the accused device," "[t]here can be no

infringement as a matter of law." *London v. Carson Pirie Scott & Co.*, 946 F.2d 1534, 1539 (Fed. Cir. 1991).

Patent eligibility under 35 U.S.C. § 101, Amazon's second ground for summary judgment, constitutes "a question of law that may involve underlying questions of fact," although "not every [patent eligibility] determination contains genuine disputes over the underlying facts material to the § 101 inquiry." *Trinity Info Media, LLC v. Covalent, Inc.*, 72 F.4th 1355, 1360 (Fed. Cir. 2023). Invalidity, including subject-matter ineligibility, constitutes an affirmative defense, and Amazon must prove any "factual questions" underpinning that defense by "clear and convincing evidence." *Microsoft Corp. v. i4i Ltd. P'ship*, 564 U.S. 96–97, 91 (2011) (interpreting 35 U.S.C. § 282(a)). [18]

On the substance of the § 101 inquiry, little can be added to Judge Ellis's exposition in the *Alice* Opinion. In short, 35 U.S.C. § 101, which defines patent-eligible subject matter, stands subject to a judicially imposed "implicit exception" that excludes "abstract ideas" from the statute's scope. *Alice Op.*, 2023 WL 7381551, at *3 (quoting *Mayo Collaborative Servs. v. Prometheus Labs., Inc.*, 566 U.S. 66, 70 (2012) and *Alice*, 573 U.S. at 216). Under the *Alice* doctrine, a court must first determine whether challenged patent claims "are directed to" an abstract idea. *Id.* (quoting *Alice*, 573 U.S. at 217). If so, the court must then "search for an inventive concept" by excising the abstract idea and considering the claims' elements "both individually and as an ordered combination to determine whether . . . the additional elements transform the nature of the claim into a patent-eligible application." *Alice*, 573 U.S. at 217–18.

---

[18]   Because "the ultimate question of patent validity is one of law," the "clear and convincing" standard of proof applies only to subsidiary questions of fact and not to the invalidity question itself. *i4i*, 564 U.S. at 96–97; *id.* at 114 (Breyer, J., concurring).

IV.   NONINFRINGEMENT[19]

Summary judgment of noninfringement must be granted in part and denied in part. Alexa

does not infringe Claim 10 of the '006 Patent or the entirety of the '327 Patent as a matter of law.

With respect to all other claims of all other Asserted Patents, disputes of fact foreclose a grant of

summary judgment.

A.   The Doctrine of Equivalents

The Court begins with a preliminary matter. In its opening brief, Amazon contends that

Alexa does not meet the literal terms of the Asserted Patents and therefore does not infringe as a

matter of law. In response, Dialect argues that Alexa does satisfy those literal terms. In addition,

at various points throughout its opposition brief, Dialect argues in the alternative that even if

Alexa does not literally infringe, summary judgment should nonetheless be denied, because a

jury could find infringement under the doctrine of equivalents.[20] (ECF No. 255 ("Dialect's

Sealed Opp.") at 16–17 ('957 Pat.), 20 ('468 Pat.), 22–25 ('039 Pat.).) In its reply brief, Amazon

contends that the Court should either strike Dialect's doctrine-of-equivalents arguments as

untimely or disregard them as insufficient. (ECF No. 277 ("Amazon's Reply") at 2.)

Amazon asserts its argument in a single paragraph of its reply brief, and Dialect has had

no opportunity to respond. Dialect, for its part, spends hardly any more time developing its

---

[19]   "[T]he specifications and inner workings of Amazon's Alexa service" constitute "highly
confidential information" that the Court has consistently allowed the parties to file under
permanent seal. (ECF No. 244 at 2.) Because the Court has no intention of disclosing Alexa's
workings to the public, this Memorandum Opinion does not summarize which facts about
Alexa's functioning are undisputed. The noninfringement portion of this Memorandum Opinion
thus presumes familiarity with the parties' briefing and the underlying technology.

[20]   Under the doctrine of equivalents, "a product or process that does not literally infringe
upon the express terms of a patent claim may nonetheless be found to infringe if there is
'equivalence' between the elements of the accused product or process and the claimed elements
of the patented invention." *Warner-Jenkinson Co. v. Hilton Davis Chem. Co.*, 520 U.S. 17, 21
(1997).

doctrine-of-equivalents theories, notwithstanding the Federal Circuit's instructions that the doctrine cannot be asserted by a plaintiff absent "particularized testimony and linking argument as to the insubstantiality of the differences between the claimed invention and the accused device." *VLSI Tech. LLC v. Intel Corp.*, 87 F.4th 1332, 1343 (Fed. Cir. 2023) (quotation omitted). The parties say little to explain their positions.

The parties' cursory mentions of their asserted doctrines leave the Court unable to meaningfully analyze their contentions without stepping outside the judicial role. It is the parties' task, not the Court's, to present the arguments that entitle them to relief. *United States v. Sineneng-Smith*, 590 U.S. 371, 375–76 (2020); *cf. Hensley ex rel. North Carolina v. Price*, 876 F.3d 573, 580 n.5 (4th Cir. 2017) (as a matter of appellate procedure, courts will not consider underdeveloped arguments). Because, with respect to the '957, '468 and '039 Patents, Dialect's theories of literal infringement present jury questions that preclude summary judgment, *see infra* §§ IV.D–F, the Court proceeds without addressing the doctrine of equivalents.

B.   The '006 Patent

Amazon's noninfringement arguments target each of the three independent claims (Claims 1, 5 and 10) asserted by Dialect. The Court will deny summary judgment with respect to Claims 1 and 5 but grant it with respect to Claim 10 and all claims dependent on Claim 10.

1.   Claim 1

Amazon argues that Alexa does not satisfy three different limitations of Claim 1. The Court addresses each one in turn.

a.   *"Recognizing"*

To begin, Amazon targets Claim 1's second element, whose text is reproduced below in relevant part:

20

> [R]ecognizing, at a speech recognition engine coupled to [a] processing device, one
> or more words or phrases contained in [an] utterance using information in one or
> more dictionary and phrase tables, wherein recognizing the one or more words or
> phrases contained in the utterance includes:
>
>   . . . .
>
>> determining an identity associated with a user that spoke the utterance based
>> on voice characteristics associated with the utterance[.]

'006 Pat. col. 25, ll. 54–65. As Amazon would have it, Alexa lacks any component identifiable

as a "speech recognition engine" that "recogniz[es] . . . words or phrases" and "determine[es] an

identity associated with a user." Amazon asserts as undisputed that only Alexa's Automated

Speech Recognition ("ASR") component functions as a "speech recognition engine," and that

Alexa's Voice ID function alone determines a user's identity based on voice characteristics.

(ECF No. 234 ("Amazon's Sealed Br.") at 9–10.) But "[b]ecause the ASR does not use Voice

ID," Amazon contends that Alexa lacks a speech recognition engine that determines user identity.

(*Id.* at 10.)

    Dialect responds that Alexa's "speech recognition engine" consists of "portions of

Amazon's ASR in combination with portions of the Voice ID service." (Dialect's Sealed Opp.

at 8.) In other words, although Alexa uses the ASR component to convert user speech into text,

Dialect avers that Alexa uses other components as part of that process as well, such that those

components constitute the '006 Patent's "speech recognition engine" when combined. Dialect

thus accuses Amazon of engaging in arbitrary line-drawing, and it disputes that Amazon's

decision to "organize[] its source code in a particular manner" should determine which lines of

code map onto which elements of the claim. (*Id.*)

    The parties, in short, do not dispute the meaning of the claim language; they dispute how

that claim language maps onto Alexa itself. As the Federal Circuit has made clear, this

constitutes a factual dispute, and the Court may grant summary judgment only if "no reasonable

factfinder" could find that Alexa meets Claim 1's requirements. *Tessera*, 646 F.3d at 1364. As the Court sees it, the relevant Alexa components consist entirely of computer code separated by electrons and bits rather than by physical space. Amazon presents no caselaw that supports treating this sort of digital distinction as conclusive. The Court finds this dispute to be factual, genuine and material, and therefore the Court will submit it to a jury.

b.   *"Parsing"*

Amazon next targets Claim 1's third element, which requires, in relevant part, "parsing, at a parser . . . , information relating to the utterance to determine a meaning associated with the utterance and a context associated with the request contained in the utterance." '006 Pat. col. 26, ll. 5–8.[21] Amazon argues that Alexa cannot meet this limitation, because the only component of Alexa that analyzes words to determine their meaning is the Natural Language Understanding unit ("NLU"), and the NLU ████████████████████████████████

████████████ " (Amazon's Sealed Br. at 11.) In other words, according to Amazon, the NLU

████████████████████████████████ rather than *a* context, as the claim language seemingly requires.

In response, Dialect again accuses Amazon of arbitrary line-drawing, and it asserts that Alexa's "parser" need not be limited to "a specific organizational structure chosen by Amazon." (Dialect's Sealed Opp. at 9.) Dialect asserts that Alexa's ████████████████████

██

████████████████████████████████ " (*Id.*) According to Dialect,

---

[21]   The Court previously construed "parser" to mean "software that analyzes a string of words" and "context" to mean "the subject matter area to which a particular user input is directed and which is used to determine the meaning of the user input." *Markman Op.*, 2024 WL 1859806, at *22, *10.

although Alexa takes multiple steps to determine the meaning and context behind a user's

utterance, the fact that these steps occur in multiple components does not avoid infringement.

The Court must note that the parties do not dispute that ██████████████████

█████████████████████████████████ Likewise, they do not dispute

that ██

██

██

██

█████████████████████████████. (Amazon's Sealed Br. ¶¶ 14–17, at 4–5; Dialect's Opp.

¶¶ 14–17, at 3.)

The parties do, however, dispute whether Alexa's ████████ constitutes part of a parser

or not. Amazon contends that the ████████ cannot be part of the parser, because the

███

████████████." (ECF No. 278 ("Amazon's Sealed Reply") at 4–5 (emphasis in original).)

But a reasonable jury applying the Court's construction of "parser" to mean "software that

analyzes a string of words" could conclude otherwise — for instance, by considering the NLU

and the ████████ to constitute the relevant "parser" in combination. Because a jury could

reasonably reject the lines that Amazon has drawn around its code and accept Dialect's

infringement theory instead, the Court finds a genuine dispute of material fact.

Amazon resists the Court's conclusion by leaning on the *Markman* Opinion. Amazon

contends that because the Court stated in the *Markman* Opinion that "[an] input can have one and

only one context," the NLU's ██████████████████ cannot infringe. (Amazon's Sealed

Br. at 11 (quoting *Markman Op.*, 2024 WL 1859806, at *9).) However, Amazon overreads the

Court's remarks. At the *Markman* stage, the Court addressed what meaning "context" carried in the claims at issue, but it was not asked and did not decide whether "context" could be singular or plural. In its original context, Amazon's quoted language merely rejects Dialect's attempts to broaden the meaning of the term "context" by using it in its "colloquial," "sum of the circumstances" sense. *Markman Op.*, 2024 WL 1859806, at *9, *7. The Court did not intend to say anything more; any language suggesting otherwise was dictum.

Contrary to Amazon's argument, longstanding Federal Circuit precedent establishes that, in the absence of "a clear intent in the claims themselves," "an indefinite article 'a' or 'an' in patent parlance carries the meaning of 'one or more.'" *Convolve, Inc. v. Compaq Comput. Corp.*, 812 F.3d 1313, 1321 (Fed. Cir. 2016) (internal quotation omitted). Without clear evidence in "the claims themselves, the specification, or the prosecution history [that] necessitate[s] a departure from the rule," *Salazar v. AT&T Mobility LLC*, 64 F.4th 1311, 1315 (Fed. Cir. 2023) — evidence that Amazon has not presented — the Court must assume that Claim 1 permits its parser to output more than one "determined context."[22] In summary, whether Alexa satisfies the "parsing" element of Claim 1 must be decided by a jury at trial.[23]

c.   "Formulating"

Finally, Amazon attacks Claim 1's fourth element, which requires "formulating, at the parser, the request contained in the utterance in accordance with a grammar used by a domain

---

[22]   This conclusion holds even though Claim 1 expressly refers to other elements using the phrase "one or more," as in "one or more words or phrases." '006 Pat. col. 25, l. 58. In *Convolve*, the Federal Circuit found itself unable to construe "a processor" as singular even though "the patentee recited other claim terms in the plural, e.g., 'output commands,' 'alter settings,' or 'input signals.'" 812 F.3d at 1321. That fact did not "compel a departure" from the Federal Circuit's "general rule." *Id.* Similar language requires the same conclusion here.

[23]   Because the Court rejects Amazon's reasoning here, the Court need not separately consider Amazon's identical arguments regarding Claims 1 and 31 of the '720 Patent. (Amazon's Br. § III.B.1, at 17–18.)

agent associated with the determined context." '006 Pat. col. 26, ll. 14–16. Amazon rests on arguments quite like those that it makes with respect to the "parsing" limitation. It contends that Claim 1 requires an utterance's context to be "determined *before* the formulating step," and that no infringement can occur because Alexa does not identify any contexts until *after* completing its "formulating" step. (Amazon's Sealed Br. at 13–14 (emphasis in original)). Dialect contests this premise and argues that the claims do not require that "the context must be determined before the formulating step." (Dialect's Sealed Opp. at 10.)

The relevant claim language merits summarization. Claim 1's third and fourth steps, in simplified terms, require that a parser (i) determine a meaning and a context and (ii) formulate a request for processing by a domain agent associated with that context. '006 Pat. col. 26, ll. 5–16.[24] Dialect's expert, Dr. H. V. Jagadish, suggests that Amazon's speechlets constitute domain agents and that their associated subject matter areas constitute contexts. For instance, the context "music" has Alexa's music speechlet as an associated domain agent. (ECF No. 234-2 ("Jagadish Op.") ¶ 405.)

The parties' arguments are not a model of clarity. The Court begins with Dialect's contentions. If Dialect argues that Alexa need not determine a single context before it performs the "formulating" step, then Dialect argues correctly. Absent clear evidence to the contrary, the Court must assume that identifying more than one "determined context" satisfies the claims. *See supra* § IV.B.1.b. If Dialect, on the other hand, argues that no context need be determined before

---

[24]   In full, the cited claim language reads as follows:

> parsing, at a parser coupled to the processing device, the one or more recognized words or phrases to determine a meaning associated with the utterance and a context associated with the request contained in the utterance, wherein the one or more recognized words or phrases are further associated with the determined context in response to the one or more recognized words or phrases satisfying the predetermined confidence level; [and]

> formulating, at the parser, the request contained in the utterance in accordance with a grammar used by a domain agent associated with the determined context[.]

Alexa performs the "formulating" step, then Dialect errs. The Court can assign no meaning to the phrase "the determined context" other than "the context that has been determined." Indeed, the Court sees no way in which a device could perform the formulating step without first determining at least one context (and thereby a domain agent and a grammar).

For its part, Amazon argues that Dr. Jagadish maps "the determined context" to Alexa's speechlets. (Amazon's Sealed Br. at 13.) The Court does not agree with that reading of Jagadish's report. Dr. Jagadish's "music" example identifies "playing music" as a context, in accordance with the Court's construction of the term to refer to a "subject matter area" rather than a particular software module. (Jagadish Op. ¶¶ 405, 408.) Dr. Jagadish opines that Alexa performs the "formulating" step when Alexa's "████████████████████████████████████ ███████████████████████████" (*Id.* ¶ 409.) A reasonable juror could favor this interpretation over Amazon's view that Jagadish maps the "formulating" step onto "████████████████████ ████████████████████████████████" (Amazon's Sealed Br. at 13 (citing Jagadish Op. ¶ 413).) This dispute presents a genuine and material question, and it should be decided by a jury.[25]

2.   Claim 5

Claim 5 of the '006 Patent expands on the "formulating" element discussed in the previous section. In Claim 5, unlike in Claim 1, "formulating the request in accordance with the grammar used by the domain agent" includes four discrete processes that must be completed by the accused device's parser. '006 Pat. col. 27, ll. 33–47. Amazon contends that Dialect cannot identify how Alexa satisfies any of these limitations. (Amazon's Br. at 14–15.) Dialect responds that the parties' competing expert opinions create material issues of fact, as Dr. Jagadish

---

[25]   For the reasons stated in this section, the Court rejects Amazon's substantively identical arguments regarding Claims 1 and 31 of the '720 Patent. (Amazon's Br. § III.B.3, at 19.)

identifies how each limitation can be found in Alexa. (Dialect's Opp. at 10–11 (citing ECF No. 255-4 ("Jagadish Reply") ¶¶ 179–83).)

Dialect contends that Alexa's "parser" cannot be limited to only the NLU, and it insists that Dr. Jagadish's rebuttal report shows that Alexa can satisfy each element. As the Court determined with respect to Claim 1, the parties genuinely dispute the scope of the parser. *See supra* § IV.B.1.b. Here, for the same reasons, the Court finds that the experts' competing reports establish a genuine and material factual dispute.

3.   Claim 10

Finally, Amazon attacks Dialect's infringement theory of Claim 10. Here, it finds greater success. Claim 10 expands on the last step of the process disclosed by Claims 1 and 5. That step requires "presenting the generated response" to the user "via the speech unit, wherein [that process] includes" the following steps:

> selecting, by the domain agent, a format template to use in presenting the generated response;
>
> selecting, by the domain agent, a personality to use in presenting the generated response;
>
> determining, by the domain agent, an order few to use [*sic*] in presenting one or more tokens contained in the generated response; and
>
> performing, by the domain agent, one or more variable substitutions and transformations on the one or more tokens contained in the generated response to vary a terminology used in presenting the generated response.

'006 Pat. col. 30, ll. 13–27. Amazon argues that Dialect fails to meet its burden of production with respect to both "selecting" steps and the "determining" step. The Court agrees with Amazon. Although Dialect identifies domain agents that perform each of these steps individually, its infringement contentions must fail, because it nowhere explains how a single domain agent (or single set of domain agents) can perform all four steps.

To illustrate the point, consider the speechlets that Dialect contends satisfy the first two limitations listed above. Dialect contends that Alexa meets the first "selecting" limitation, because ███████████████████████████████████████. Thus, Dialect says, these "███████████████████████████████████." (Dialect's Sealed Opp. at 11–12.) In support of this view, Dr. Jagadish identifies speechlets like the weather speechlet and the shopping speechlet. (Jagadish Reply ¶ 193.) The Court assumes without deciding that these speechlets satisfy the first "selecting" limitation. For the "personality" limitation, however, Dialect identifies completely different domain agents. At this step, Dialect points to "third party skills" like "Hey Disney!" that allow Alexa to respond to queries as one of several cartoon characters. (Jagadish Op. ¶ 464.) The Court again assumes without deciding that "Hey Disney!" constitutes a "domain agent" in its own right.

This inconsistency renders Dialect's infringement theory legally insufficient. Claim 10 specifies that each step (i.e., "selecting . . . a format template" and "selecting . . . a personality") must be performed by "the domain agent." '006 Pat. col. 30, ll. 16–19. Typically, "[t]he use of the definite article, 'the,' means that the phrase [introduced using the definite article] refers back to earlier language." *ABS Glob., Inc. v. Cytonome/St, LLC*, 84 F.4th 1034, 1040 (Fed. Cir. 2023). "The reference-back 'the' language takes its meaning from the meaning of [its] antecedent." *Id.* (collecting cases). Thus, when Claim 10 states that "presenting the generated response includes" four different steps connected using the word "and," and each of them referring to "the domain agent," the only reasonable reading of the Claim is that the same domain agent (or domain agents) must perform each of the four steps recited. *See Lucent Techs., Inc. v. Gateway, Inc.*, 525 F.3d 1200, 1214 (Fed. Cir. 2008) (the word "including" typically means that the listed steps "are

essential"). Amazon therefore hits the mark when it argues that Dialect must identify at least one domain agent that performs all four steps. (Amazon's Reply at 8.) Dialect does not do so.

Dialect has no evidence that any domain agent in Alexa can perform each of the steps of Claim 10's "presenting" element. For that reason, Dialect has not met its burden of production. Summary judgment of noninfringement must be granted to Amazon with respect to Claim 10 of the '006 Patent.[26]

C.   The '720 Patent

Amazon addresses Claims 1 and 31 of the '720 Patent together. Because the two claims share substantially similar language, *see supra* n.9, the Court does the same. Amazon makes four distinct arguments for noninfringement. Because the Court has already rejected two of those arguments, *see supra* nn.23, 25, the Court discusses only the remaining two.

1.   "Selecting"

The Court begins with element (a)(2)(B). *See supra* n.8. That element recites "a parser that interprets [] recognized words or phrases . . . by . . . selecting at least one of [a] plurality of domain agents based on [a] determined context." '720 Pat. col. 32, ll. 51–59 (Claim 1); *see id.* at col. 35, ll. 42–44 (Claim 31) (similar). As it argued with respect to the '006 Patent, Amazon submits that only Alexa's NLU can be a "parser," and thus Dialect's identified corresponding structure (Alexa's NLU combined with portions of its Dynamic Routing component) could not possibly infringe. (Amazon's Sealed Br. at 18.) The Court can dispatch this argument quickly; it proves to be no more than another variation on Amazon's "line-drawing" theme. As before, the parties genuinely dispute how Alexa's components correspond to the text of the claims, and as

---

[26]   The Court must also grant summary judgment of noninfringement on dependent Claim 11 of the '006 Patent, which incorporates Claim 10's terms. '006 Pat. col. 30, ll. 28–35.

29

before, the Court finds this dispute sufficiently genuine to defeat summary judgment. *See supra* §§ IV.B.1.a., IV.B.1.b.

Amazon also argues that neither Alexa's NLU nor its Dynamic Routing component select "at least one of [a] plurality of domain agents based on the determined context." (Amazon's Sealed Br. at 18.) This must be so, per Amazon, because "Dr. Jagadish maps the same functionality of determining or selecting the 'most likely speechlets and skills'" to two different limitations — "determining a context" and "selecting . . . domain agents based on the determined context" — and "[t]he same functionality cannot meet both requirements." (*Id.* (quoting Jagadish Op. ¶¶ 349, 352).)

Dialect responds by clarifying its theory of infringement. In Dialect's view, Alexa selects domain agents — i.e., speechlets or skills — based on "███████████████████████████ ██████" — i.e., Alexa's ████████████████████. (ECF No. 357 ("Dialect's Sealed Suppl. Br.") at 2.) The Court finds this factual argument sufficient to satisfy Dialect's burden of production. If Alexa's ██████████ constitutes a set of contexts, as the Court has already determined it plausibly might, *see supra* §§ IV.B.1.b, IV.B.1.c, then a rational jury could agree that Alexa selects domain agents on the basis of those contexts. Therefore, a genuine dispute of fact exists as to this element.[27]

---

[27]   In its opposition and in its supplemental brief, Dialect also argues that Alexa may select a domain agent "based on a context that is later identified as the *determined* context." (Dialect's Sealed Suppl. Br. at 2 (emphasis in original); *see* Dialect's Sealed Opp. at 13 (similar).) The Court rejects that argument. The claims unambiguously require that domain agents be selected "*based on* the determined context." '720 Pat. col. 32, l. 59 (emphasis added). Dialect survives at this stage not because "Alexa eventually selects the most appropriate context for an utterance" (Dialect's Sealed Suppl. Br. at 2), but rather because Alexa's ██████████ plausibly constitutes a set of determined contexts, and "the determined context" must, absent some contrary indication, be interpreted to mean "one or more contexts." *See supra* § IV.B.1.b.

2.   "Receive, Process, and Respond"

Finally, Amazon addresses two similar portions of Claims 1 and 31. There, the '720

Patent recites "a speech unit . . . [that] converts [a] received natural language speech utterance

into an electronic signal." '720 Pat. col. 32, ll. 32–36 (Claim 1); *id.* at col. 35, ll. 25–28

(Claim 31). Claim 1 then recites "a natural language speech processing system . . . [that]

receives, processes, and responds to the electronic signal using data received from a plurality of

domain agents." *Id.* at col. 32, ll. 37–41. For its part, Claim 31 recites "a speech recognition

engine . . . [that] uses at least data received from a plurality of domain agents to recognize []

words or phrases." *Id.* at col. 35, ll. 29–34. Amazon argues that Dialect can identify no "data

received from a plurality of domain agents" that Alexa uses in the manner that Claims 1 and 31

require. (Amazon's Br. at 19–20.)

Dialect argues in response that "Alexa receives data from multiple domain agents and

uses that information to respond to a user query." (Dialect's Sealed Opp. at 14.) To support this

argument, Dialect cites to Dr. Jagadish's reply report, which contends that a combination of

Alexa components "including without limitation the ████████, HypRank, Dynamic Routing,

ASR, and NLU" collectively constitute a natural language speech processing system that uses

data from domain agents to "receive, process, and respond to electronic signal." (Jagadish Reply

¶ 110.) In Dialect's view, no party disputes that "Alexa uses speechlets to respond to user

queries," and because those speechlets qualify as "domain agents," Alexa thus receives data from

domain agents to respond to the relevant electronic signal. (Dialect's Sealed Opp. at 14.)

Dialect also argues that Claim 1 does not require using data received from a plurality of domain

agents to receive or process the electronic signal in question, such that responding to the signal

stands sufficient to satisfy the claim limitation. (Dialect's Opp. at 14.) In reply, Amazon

highlights that Dialect fails to show that any component in Alexa "receive[s] data from a

plurality of domain agents" to receive, process "*and*" respond to an electronic signal for a single natural language speech utterance. (Amazon's Reply at 11 (emphasis in original).)

Three matters merit clarification. First, it appears that, as Dialect contends, Amazon's argument applies only to Claim 1 of the '720 Patent. Amazon does not argue that Alexa fails to "use[] at least data received from a plurality of domain agents to recognize [] words or phrases." '720 Pat. col. 35, ll. 32–34. Second, Amazon's insistence that "the ASR or NLU" must constitute Alexa's "natural language speech processing system" repeats line-drawing arguments that the Court cannot adjudicate at the summary judgment stage. (Amazon's Sealed Br. at 19); *see supra* § IV.B.1.a. And finally, although Amazon argues that "data from a plurality of domain agents" must be used to respond to "a *single* 'natural language speech utterance'" (Amazon's Reply at 11 (emphasis in original)), Amazon grounds its contention entirely in Claim 1's use of indefinite articles, i.e., in Claim 1's recitation of "*a* natural language speech utterance" and "*an* electronic signal." '720 Pat. col. 32, ll. 33–36 (emphasis added). However, as the Court has already explained, "in the absence of 'a clear intent in the claims themselves,' 'an indefinite article "a" or "an" in patent parlance carries the meaning of "one or more."" *See supra* § IV.B.1.b (quoting *Convolve, Inc.*, 812 F.3d at 1321). Amazon presents no evidence that the patentee intended to require more than one domain agent to respond to only one electronic signal, so the Court must reject this argument.

The Court proceeds to evaluate the parties' remaining arguments regarding Claim 1. Because these arguments sound in claim construction, the Court can resolve them as a matter of law, employing the principles that the Court relied on at the *Markman* stage of this proceeding.[28]

---

[28]   The Court's *Markman* Opinion thoroughly summarized the basic rules of claim construction. 2024 WL 1859806, at *2–3.

The '720 Patent requires a "natural language speech processing system [that] receives, processes,

and responds to the electronic signal using data received from a plurality of domain agents."

'720 Pat. col. 32, ll. 38–41. Amazon insists that each of the three steps of that process —

receiving, processing and responding — must be performed "using data received from a plurality

of domain agents." In effect, Amazon would have the Court read the quoted claim language to

require a system that "receives [the electronic signal using data received from a plurality of

domain agents], processes [the electronic signal using data received from a plurality of domain

agents], and responds to the electronic signal using data received from a plurality of domain

agents."

Amazon's reading of Claim 1 may be "grammatically permissible," but "reviewing text

in context" reveals it to be semantically implausible. *Pulsifer v. United States*, 601 U.S. 124, 133

(2024).[29] The remainder of Claim 1 specifies how a plurality of domain agents can be used to

process and respond to the electronic signal in question. For instance, a part of the natural

language speech processing system must recognize words or phrases encoded in the electronic

signal "using at least the data received from the plurality of domain agents." '720 Pat. col. 32,

ll. 43–46. Later, the claimed system must interpret the words or phrases extracted from the

electronic signal by "determining a context" for the utterance being analyzed and "selecting at

least one of the plurality of domain agents based on the determined context." *Id.* at ll. 51–59.

That "selected domain agent" then produces "a response" to the electronic signal and "format[s]

the response for presentation to the user." *Id.* at col. 33, ll. 3–6. But neither the claim, the

---

[29]   *Pulsifer* has some parallels to the question presented here. In that case, the Supreme
Court asked whether a statute's text required "that a defendant 'does not have (A, B, and C),' . . .
[or] that he 'does not have A, does not have B, and does not have C.'" 601 U.S. at 136. Because
grammar proved ambiguous, the Supreme Court resorted to context. Confronted with similar
text, this Court does the same.

specification, nor Amazon suggest how "data received from a plurality of domain agents" can be used to receive an electronic signal transmitted from "a speech unit" to "a natural language processing system." *Id.* at col. 32 ll. 32–41. The Court finds it difficult to imagine how data would aid that process.

A far more natural reading of the claim language would be, as Dialect suggests, that the phrase "using data received from a plurality of domain agents" modifies the entire phrase "receives, processes, and responds to the electronic signal." The Court finds a genuine dispute of fact regarding whether a "natural language speech processing system" within Alexa uses "data received from a plurality of domain agents" as part of that overall process. Consequently, Amazon's argument does not warrant summary judgment.

D.   The '957 Patent

Next, the Court turns to the '957 Patent. Amazon presents one argument for noninfringement of Claim 1, the only independent claim of the '957 Patent that Dialect asserts. Claim 1, in relevant part, requires the recited "system" to generate a context stack containing "a plurality of context entries," receive a spoken message, determine "one or more words" contained in that message, and then "identify, from among the plurality of context entries, one or more context entries that correspond to the one or more words . . . [by] comparing the plurality of context entries to the one or more words." '957 Pat. col. 39, ll. 38–58. Amazon asserts that Alexa performs no such comparison. (Amazon's Br. at 20.)

Dialect responds by accusing Amazon of yet another "flawed attempt at arbitrary box drawing." (Dialect's Sealed Opp. at 15.) As Dialect would have it, Alexa contains a function called ███████████████████████████████████

████████████████████████████████████," which Dialect maps onto the '957 Patent's context entries. Meanwhile, "Alexa's ASR component captures the user's spoken words, [and]

34

Alexa's NLU component reformats [those] words into ███████████." (*Id.*) An Alexa component called ███████████ then ██████████████████████████████ ████████████████████████████ thus performing the comparison that Claim 1 requires.

Amazon responds by accusing Dialect of mischaracterizing both its expert's testimony and Alexa itself. As Amazon would have it, "Dr. Jagadish never opined that the NLU's interpretations" — what Dialect calls "reformatted words" — constitute words in themselves. (Amazon's Sealed Reply at 12.) And Amazon emphasizes that the '957 Patent operates on words and phrases only, not some equivalent. (*Id.*)

As the Court sees it, the parties present a dispute of fact. Alexa cannot operate on words and phrases directly, if only because words and phrases constitute a part of natural language rather than a data type that a computer can understand. Dialect asserts that Alexa represents words and phrases using a specific format ██████████████████████████████ ██████████████████████. A jury must decide whether that theory corresponds to the undisputed meaning of Claim 1's text. The Court, in other words, finds Alexa's infringement of the '957 Patent to be genuinely disputed.[30]

E.   The '468 Patent

Amazon presents four reasons why Alexa does not infringe the '468 Patent. One of those reasons reasserts arguments that the Court already rejected regarding the '957 Patent. *See supra* n.30. To the extent that Amazon renews that argument here, the Court does not address it.

1.   "Transcribing"

Claim 19 of the '468 Patent, the only independent claim asserted, concerns itself with the "processing" of "multi-modal natural language inputs." '468 Pat. col. 40, ll. 37–38. The method

---

[30]   For the same reasons stated above, the Court rejects Amazon's identical argument regarding Claim 19 of the '468 Patent. (Amazon's Br. § III.D.4, at 23.)

that Claim 19 describes requires receiving a multi-modal input — that is, an input that has both speech and non-speech components — and having a "transcription module . . . transcribe[] the non-speech input to create a non-speech-based transcription" separately from transcription of the speech input. *Id.* at ll. 42–45. Amazon contends that Alexa does not "transcribe" any non-speech inputs and therefore cannot infringe.

As Amazon would have it, Alexa receives two different kinds of inputs: text inputs and touch inputs. (Amazon's Br. at 21.) Text inputs need not be transcribed at all, because text inputs can be used "as is" without invoking any module; and touch inputs "require[] no transcription or interpretation" at all. (Amazon's Sealed Br. at 21–22.) Dialect counters by asserting that Amazon's argument assumes "that a transcription of user input must be a 'word-for-word' verbatim copy," and instead proposes that a person of ordinary skill "would understand that 'transcription' . . . means . . . a knowledge-based transcription." (Dialect's Sealed Opp. at 17–18.)

Amazon supports its argument by resting on a portion of the '468 Patent's specification. There, when describing "one embodiment of the invention," the specification states that "one or more speech recognition engines" may "transcribe utterances to textual messages." '468 Pat. col. 23, ll. 56–59. Then, the specification discloses that in a different "embodiment," "data that is entered in a text format may be merged with data that is transcribed to a textual format from the utterance." *Id.* at ll. 62–64. Thus, Amazon argues, the patent itself contemplates that textual inputs would not be "transcribed."

Amazon's reliance on the specification falls flat. Although the specification refers to textual data being "entered," this does not imply that textual data cannot also be "transcribed" within the meaning of Claim 19. To the contrary, the specification strongly supports an inference

that transcription of text inputs is possible. The '468 Patent's specification repeatedly recites "keypads" and "keyboards" as possible parts of the claimed invention. *See, e.g.*, '468 Pat. col. 16, ll. 18–19 ("keypads **14** for receiving textual data input"), col. 19 ll. 1–4 ("users may interact with the mobile device **36** through . . . the keypad **74** or keyboard"). The patent thus contemplates text constituting a form of "non-speech input." But the Court can find no reference in the claims to a non-speech input being "entered." The invention only ever "transcribes" non-speech inputs. The Court does not find it plausible that the patentee intended not to claim something that was so repeatedly and specifically described as within the scope of the '468 Patent. Indeed, the specification supports an inference that this embodiment was *included* in the claimed invention rather than excluded from it. The Court does not agree that, if text was "entered" in a device, then that text has necessarily not been "transcribed" and thus lies outside the scope of the claims. The two categories are not mutually exclusive.

To summarize, both parties find plenty in Alexa to ground their respective arguments for and against infringement. The parties genuinely dispute whether the way that Alexa processes text constitutes "transcription." For that reason, Amazon cannot be granted summary judgment.

2. "Identifying"

Next, Amazon asserts that Alexa does not "identify[] the user that provided the multi-modal input" and then "creat[e] a speech-based transcription . . . using . . . a semantic knowledge-based model . . . [that] includes a personalized cognitive model derived from one or more prior interactions between the identified user and the conversational voice user interface." '468 Pat. col. 40, ll. 46–53. Dialect disagrees. Dialect contends that Alexa uses "███████████ ██." (Dialect's Sealed Opp. at 18–19.) Dialect also claims that Alexa creates a "personalized cognitive model" using modules like "█████████████████████████████████████." (*Id.*) In its reply, Amazon

does not address Dialect's references to ███████████████████████, and it asserts

that ███████████████ do not identify the user. (Amazon's Sealed Reply at 13–14.)

Amazon's argument hinges on its assertion that Claim 19 requires "using models

associated with the identified user." (Amazon's Sealed Br. at 22 (emphasis removed).) That

assertion misrepresents the claim, which instead requires using a model "derived from one or

more prior interactions between the identified user and the conversational voice user interface."

'468 Pat. col. 40, ll. 51–53. The parties dispute that element, since the parties disagree, for

example, whether ████████████████████████████████████

███████████████████" (Dialect's Sealed Opp. at 19) or instead merely "██████

███████████████████████████████. (Amazon's Sealed Reply

at 14.) Because the nature of Alexa's inner workings remains thoroughly disputed, the Court

finds a genuine dispute of fact regarding whether Alexa relies on a model "derived from one or

more prior interactions with the identified user."

3. "Merging"

Third, Amazon looks to the fourth element of Claim 19's recited method, which requires

"merging the speech-based transcription and the non-speech-based transcription to create a

merged transcription." '468 Pat. col. 40, ll. 59–61. Amazon asserts that "████████████

███████████████████████████████████. (Amazon's Sealed

Br. at 22.) Thus, it says, Alexa never combines the speech and non-speech inputs into one

transcription. (*Id.*)

Dialect does not contest that Alexa ██████████████████████████

███████████████████. Instead, it argues that "Alexa can operate on a speech input

and a non-speech input in a multiturn user interaction." (Dialect's Sealed Opp. at 20.) In

Dialect's view, Alexa's "███████████████████" constitutes a set of transcriptions. (*Id.*)

When a user gives Alexa a non-speech input, Alexa generates ████████████████████

███

██ (*Id.*) ███████

███

████████████████████████████████

The parties, in short, dispute whether the inputs to ████████████ may best be

described as ████████████ (Dialect's Sealed Opp. at 20), ████████████ (Amazon's Sealed

Br. at 23), or something else; they likewise dispute whether ████████████ function can

fairly be described as "merging transcriptions." The Court will let both parties make these

factual arguments before the jury.

The Court finds that each of Amazon's three foregoing arguments regarding the '468

Patent present questions best put to a jury. Because the Court has already rejected Amazon's

fourth argument, *see supra* n.30, the Court will not grant summary judgment of noninfringement.

F.   The '039 Patent

Amazon's argument with respect to the '039 Patent begins by incorporating arguments

that Amazon made against infringement of the '468 Patent. (Amazon's Br. at 23–24.) With

respect to the '468 Patent, the Court rejected both of those arguments, namely, that Alexa neither

transcribes nor merges non-speech inputs as a matter of law. For the same reasons, the Court

rejects them again now. *See supra* §§ IV.E.1, IV.E.3.

1.   "Accessing a Plurality"

In its first new argument, Amazon attacks the sixth element of the '039 Patent's Claim

13. That element recites, as part of its claimed method, "accessing a plurality of domain agents

that are associated with the context description grammar." '039 Pat. col. 30, ll. 50–51. Amazon

asserts that Alexa never accesses "a plurality" of domain agents — instead, ████████████████

███.” (Amazon's Sealed Br.

at 24); *see SIMO Holdings Inc. v. Hong Kong uCloudlink Network Tech. Ltd.*, 983 F.3d 1367,

1377 (Fed. Cir. 2021) (stating that "[t]he phrase 'a plurality of' means 'at least two of'" and

collecting cases).

Dialect proposes two cognizable ways that Alexa can access a plurality of domain agents.

First, ██

██" (Dialect's Sealed

Opp. at 23.) ███

██. (*Id.* at 24.) Second, Dialect

says, ██

████████, thereby satisfying the element in question. (*Id.*)

Amazon does not rebut Dialect's contentions. Amazon merely insists in conclusory

fashion that ██

█████████████████████████████████. (*See id.* ("██████████████████████████

███████████████████████████████████████.") (quoting deposition

testimony).) But a jury could reasonably conclude that ███████████████████████████

███████████████████████████████████████, just as it could rationally find that

███████████████████████████████████████ in the relevant sense. Amazon's

contrary argument fails.

2.   "Comparing"

Claim 13 requires generating a query by merging speech-based and non-speech based

"textual messages," "searching the query for text combinations," and then "comparing the text

combinations to entries in a context description grammar." '039 Pat. col. 30, ll. 42–49. Dialect

identifies Alexa's ██████████████████████████████ as constituting its "context

description grammar." Amazon argues that this conclusion cannot be correct. (Amazon's Sealed Br. at 24.)

Amazon relies on the Court's construction of "context description grammar" at the *Markman* stage. There, the Court construed that term to mean "a data structure containing entries constituting or referencing sets of rules, wherein each of those sets describes the structure of natural language in a particular context." (ECF No. 213 (*Markman* Order) at 2.) Amazon contends that its ██ ██████████████████" (Amazon's Sealed Br. at 25.)

Dialect responds by referring to the Court's holding that the "rules" referenced by the Court's construction "may be probabilistic in nature." *Markman Op.*, 2024 WL 1859806, at *14 n.29 (cited by Dialect's Opp. at 25). Dialect says that ████████████████████████████ ██." (Dialect's Sealed Opp. at 25.) Amazon contends, on the other hand, that "████████████████████████████████ ████████████████████." (Amazon's Sealed Reply at 16.)

The parties' dispute, once more, is predicated on different understandings of the facts. On summary judgment, the Court may not weigh the evidence for and against Amazon's preferred view of Alexa's function. The parties genuinely dispute a material fact, namely, whether Amazon's ████████████ matches the Court's construction of "context description grammar." This dispute, like many others, must be resolved at trial.

G.   The '327 Patent

Finally, the Court comes to the '327 Patent. Amazon argues that Alexa does not display several features, each of which is recited in independent Claims 1 and 14 and incorporated by reference into dependent Claims 5, 6, 18 and 19. Evaluating all the undisputed facts presented,

and drawing all reasonable inferences in Dialect's favor, the Court concludes that Alexa does not infringe the '327 Patent as a matter of law.

Both independent claims asserted recite "a speech coder that uses . . . variable rate sampling to compress and digitize [an] input speech signal." '327 Pat. col. 26, ll. 17–20 (Claim 1).[31] As relevant here, Amazon argues that Alexa contains no such speech coder, because the component that Dialect identifies as a "speech coder" (Alexa's Opus encoder, or simply "Opus") does not digitize speech inputs. Instead, the parties do not dispute that Alexa digitizes speech inputs at its microphone, well before those inputs reach Opus. (Amazon's Sealed Br. at 26–27.)

Dialect argues that the claims at issue do not require that the input to the speech coder must be analog rather than digital.[32] (Dialect's Sealed Opp. at 27.) In support of this view, Dialect points to the '327 Patent's specification. On Amazon's reading of the claims, Dialect says, the input to the "speech coder" must be analog, contradicting the specification's statement that "[s]peech received at the microphones may . . . be processed with analog or digital filters," after which "the system may use variable rate sampling to maximize the fidelity of the encoded speech." (*Id.* (quoting '327 Pat. col. 7, ll. 43–49).) Thus, Dialect contends, Amazon's interpretation would impermissibly "exclude embodiments that use a digital filter." (*Id.*)

---

[31]   In an equivalent limitation, Claim 14 recites "a speech coder . . . configured to . . . use . . . variable rate sampling to compress and digitize [an] input speech signal." '327 Pat. col. 27, ll. 38–44.

[32]   Dialect also mentions that, in its view, Amazon surprised Dialect with this argument by asserting it "[f]or the first time" on summary judgment. (Dialect's Opp. at 26 & n.1.) Dialect, however, does not ask the Court to disregard Amazon's argument, nor does it cite any authority that would support such a step.

Dialect's argument presents a claim construction question, because Dialect does not dispute that Alexa digitizes speech inputs at its microphone rather than its speech coder. The Court can thus proceed by relying on its now-familiar *Markman* toolkit. *See Markman Op.*, 2024 WL 1859806, at *2–3 (discussing the applicable principles and cases). Applying that toolkit, the Court finds that the input speech signal recited in Claims 1 and 14 must be digitized at the speech coder and not sooner, leaving Dialect with a legally insufficient theory of infringement.

The Court begins with "the language of the claims themselves." *Trs. of Columbia Univ. v. Symantec Corp.*, 811 F.3d 1359, 1362 (Fed. Cir. 2016). The independent claims at issue — Claims 1 and 14 — both strongly suggest that digitization must occur at the speech coder and not earlier. Claim 1's text paints the picture well. Claim 1's method involves "receiving a natural language utterance at a microphone array" that alters "an input speech signal corresponding to the natural language utterance." '327 Pat. col. 25, ll. 61–67. The next steps of the claimed method each act on that "input speech signal," first "comparing" it to environmental noise, then "passing" it to an "adaptive filter," then "suppressing" noise within it. *Id.* at col. 26, ll. 1–15. The next step, however, involves "sending the input speech signal" to the speech coder, which uses various techniques "to compress and digitize the input speech signal." *Id.* at ll. 16–20. After that, the final step of the method repeatedly refers to a "digitized input speech signal," not, as before, an "input speech signal":

> transmitting the *digitized* input speech signal from a buffer in the speech coder to the speech recognition engine, wherein the speech coder transmits the *digitized* input speech signal to the speech recognition engine at a rate that depends on available bandwidth between the speech coder and the speech recognition engine.

Id. at ll. 25–30 (emphasis added).[33] The inference from this language could not be clearer:

Before the input speech signal reaches the speech coder, the signal is not digitized. The speech

coder digitizes the signal, after which it becomes a "digitized input speech signal." If all non-

digital signals must be analog, then the claim language, taken alone, requires the result that

Dialect so strongly resists.

Dialect does not offer an alternative interpretation of the claim text. Instead, it rests

exclusively on one passage from the specification and attempts to use that passage to trump the

language of the claim that it asserts. Dialect insists that, since the '327 Patent's specification

states that "[s]peech received at the microphones [of the invention] may then be processed with

analog or digital filters," the cited claim language must permit digitization at the microphone.

'327 Pat. col. 7, ll. 43–45.

The Court finds Dialect's argument unconvincing. Concededly, the invention described

in the specification appears to allow digital filtering. And it is true that claim language should be

read "in view of the specification." *Phillips v. AWH Corp.*, 415 F.3d 1303, 1315 (Fed. Cir. 2005)

(en banc). But it remains axiomatic in patent law that the claims — not the specification —

"define the metes and bounds of the patentee's invention." *Kara Tech. Inc. v. Stamps.com Inc.*,

582 F.3d 1341, 1347 (Fed. Cir. 2009). The specification cannot expand the scope of the claimed

invention any more than it can shrink it. As a predecessor to the Federal Circuit put it in an oft-

quoted maxim, "[c]ourts can neither broaden nor narrow the claims to give the patentee

something different than what he has set forth." *Autogiro Co. of Am. v. United States*, 384 F.2d

391, 396 (Ct. Cl. 1967).

---

[33]   Claim 14 features all the same patterns. That claim's "speech coder" also uses "variable
rate sampling to compress and digitize the input speech signal," after which it becomes a
"digitized input speech signal." '327 Pat. col. 27, ll. 38–54.

Moreover, the Court cannot infer that the claims of the '327 Patent monopolize the full

scope of the invention disclosed in that patent's specification. This is so, because Dialect's

chosen portion of the specification did not originate with the '327 Patent and is far from unique

to it. Recall that the '327 Patent hails from a large family and resulted from a division of the

'570 Patent, which was itself a division of the '209 Patent. *See supra* § II.F. Recall, in addition,

that the '006 Patent continues the '209 Patent. This extended patent family includes several

other members,[34] but the three relations listed here constitute a representative cross-sample. The

crucial piece of the specification that Dialect cites can be found verbatim in each of these three.

'006 Pat. col. 7, ll. 30–46; '209 Pat. col. 7, ll. 23–39; '570 Pat. col. 7, ll. 28–44. In other words,

the specification that Dialect cites was written to describe an invention disclosed in many patents

and subject to a multitude of claims. This one-size-fits-all paragraph, repeated time after time

across various patents claiming any and all aspects of the disclosed invention, does not create a

sufficiently strong inference to allow the Court to rewrite the '327 Patent's claims. Where, as

here, multiple patents lay claim to a single invention, a patentee cannot assume without

explanation that the claims of one patent capture the full scope of the invention disclosed.

In summary, the Court concludes that Alexa does not infringe the '327 Patent as a matter

of law, because Alexa digitizes speech inputs at its microphone, and the '327 Patent requires

speech inputs to be digitized only at a speech coder. Dialect's contrary argument relies on the

'327 Patent's specification to impermissibly broaden the scope of its claims. Since the Court's

---

[34]   In addition to the patents discussed by the Court, the '209 Patent family also includes U.S. Pat. Nos. 7,502,738 (a division of the '209 Patent); 8,112,275 (a division of the '570 Patent); 8,155,962 (a continuation of the '570 Patent); 8,731,929 (a continuation of the '738 Patent); and 9,734,825 (a continuation of the '929 Patent). Suffice it to say, the family tree is complex.

conclusion rests on matters of law rather than disputes of fact, the Court will grant Amazon

summary judgment on this portion of Dialect's amended complaint.

V.   SUBJECT-MATTER ELIGIBILITY

The Court now turns to Amazon's *Alice* arguments. Last November, Judge Ellis deferred

decision on Amazon's *Alice* challenge, because he believed that claim construction would

usefully illuminate the subject-matter issues in this case. *Alice Op.*, 2023 WL 7381551, at *6.

Judge Ellis was correct. Although the parties do little more than repeat the contents of their

motion-to-dismiss briefs, the Court's claim constructions nevertheless shed meaningful light on

Alice Step One. Because close analysis reveals no abstraction, each of the Asserted Patents

passes muster under *Alice* without resorting to Step Two. Accordingly, the Court will deny

summary judgment on this ground.

A.   Review of the Relevant Law

1.   The Standard

The *Alice* doctrine's Step One requires the Court to "determine whether the claims at

issue are directed to [an abstract idea]." *Alice Corp. v. CLS Bank Int'l*, 573 U.S. 208, 217 (2014).

This is easier said than done. Patents, unlike other forms of intellectual property, are peculiarly

the domain of "the idea itself." *Mazer v. Stein*, 347 U.S. 201, 217 (1954). As Judge Ellis

remarked at the motion to dismiss stage, "almost any patent can be addressed to an abstract idea,

taken at the highest level of generality." (ECF No. 80 (Hr'g Tr.) at 37:15–20.) Thus, as it must

do with respect to the *Alice* test as a whole, the Court must "tread carefully in construing" Step

One lest it "swallow all of patent law." *Alice*, 573 U.S. at 217.

To properly apply Step One, the Court must consider the claims at issue "in their entirety

to determine . . . their character as a whole." *McRO, Inc. v. Bandai Namco Games Am., Inc.*, 837

F.3d 1299, 1312 (Fed. Cir. 2016). This requires "determin[ing] whether the claims 'focus on a

specific means or method that improves the relevant technology' or are 'directed to a result or effect that itself is the abstract idea and merely invoke generic processes and machinery.'" *Apple, Inc. v. Ameranth, Inc.*, 842 F.3d 1229, 1241 (Fed. Cir. 2016) (quoting *McRO*, 837 F.3d at 1314). Thus, a common approach to Step One counsels "examining the 'focus of the claimed advance over the prior art.'" *AI Visualize, Inc. v. Nuance Commc'ns, Inc.*, 97 F.4th 1371, 1378 (Fed. Cir. 2024). Put differently, when a patent claims "computer-related technology, . . . patent claims may be non-abstract at *Alice* [S]tep [O]ne if the focus of the claimed advance is on an improvement in computer technologies rather than the mere use of computers." *Id.*

At bottom, in the absence of any "single, succinct, usable definition or test" describing "what an 'abstract idea' encompasses," the Court must apply "the classic common law methodology" and "examine earlier cases in which a similar or parallel descriptive nature can be seen." *Amdocs (Isr.) Ltd. v. Openet Telecom, Inc.*, 841 F.3d 1288, 1294 (Fed. Cir. 2016). The Court, guided by the parties' submissions, undertakes that Step One study below. If a claim proves valid under *Alice* Step One, the Court need not proceed to *Alice* Step Two. *Alice*, 573 U.S. at 217, 221.

2.   The Parties' Arguments

Amazon contends that all asserted claims of the '006 Patent, '720 Patent, and '957 Patent "are directed to the abstract idea of understanding and responding [to] a spoken request." (Amazon's Br. at 30; *see id.* at 35 (asserting that the '720 and '957 Patents display "the same abstract idea as the 006 patent.").) Amazon presses this point by emphasizing that "the claims recite no technology to achieve [their] aspirational results," nor do they present "a technological solution." (*Id.* at 30.) And Amazon believes that the same conclusion is bolstered by the fact that "the patent claims a process that can be carried out by a human." (*Id.* at 31.) Amazon assigns the '468 and '039 Patents a similar abstract idea — that of "understanding and

responding to a multimodal request." (*Id.* at 39.) In analogous fashion, Amazon charges these two patents with abstractness, because they "do not recite any technology for using context to respond to a multimodal request" and thus "do not describe how to achieve this aspirational result in a non-abstract way." (*Id.* at 38.)

Dialect does not contest that "understanding and responding to a spoken request" and "understanding and responding to a multimodal request" both constitute abstract ideas, but Dialect argues that the Asserted Patents do not claim them. Instead, Dialect says, the Asserted Patents "claim specific software components that improve the natural language understanding functionality of prior art computer systems." (Dialect's Opp. at 31.) Because the Asserted Patents "provid[e] a specific set of software architectures," says Dialect, they should not be deemed abstract. (*Id.* at 32.) Dialect claims that, by arguing to the contrary, Amazon "violates clear precedent that cautions courts not to overgeneralize the claims." (*Id.* at 34 (citing *McRO*, 837 F.3d at 1313).)

The parties' unhelpful briefing illuminates the virtues of the common-law method that the Federal Circuit endorsed in *Amdocs*. Amazon glosses over the claims and insists, in conclusory fashion, that the Asserted Patents recite results without providing answers. Dialect parrots the claims' language and insists, in an equally conclusory way, that the Asserted Patents recite specific solutions. Without discussing "earlier cases in which a similar or parallel descriptive nature can be seen," as Judge Ellis did in the *Alice* Opinion, the Court sees no way to resolve the parties' disagreement. *Amdocs*, 841 F.3d at 1294.[35]

---

[35]   Judge Ellis synthesized Federal Circuit case law to derive three relevant forms of abstraction: (1) claims reciting a process that uses computers as a tool instead of providing an improvement to computers' function; (2) claims "reciting 'mental processes' such as collecting data, analyzing or comprehending it, and then reporting the results"; and (3) claims reciting results without disclosing specific means for achieving those results. *Alice Op.*, 2023 WL

The Court proceeds by reviewing a selected subset of the cases that the parties' briefs cite. Once all of them have been described, a picture of the doctrine begins to emerge.

3.   The Parties' Cases

Amazon first cites *Affinity Labs of Texas, LLC v. Amazon.com, Inc.*, 838 F.3d 1266 (Fed. Cir. 2016). There, the Federal Circuit addressed claims concerning "targeted advertising." *Id.* at 1267. As summarized by the court, a representative claim was "directed to a network-based media system with a customized user interface, in which the system delivers streaming content from a network-based resource upon demand to a handheld wireless electronic device having a graphical user interface." *Id.* at 1268.[36] The court found this claim abstract at *Alice* Step One. It reasoned that the claim's recitation of "concrete, tangible components" did not place it beyond the reach of the *Alice* doctrine. *Id.* at 1269 (quoting *In re TLI Commc'ns LLC Pat. Litig.*, 823 F.3d 607, 611 (Fed. Cir. 2016)). Then, it rejected the plaintiff's contention that the invention embodied in the claim was concrete. As the *Affinity Labs* court reasoned, the claims did "no more than describe a desired function or outcome, without providing any limiting detail that

---

7381551, at *4 (collecting cases). That analysis proves useful to the Court, but it does not eliminate the need for case-by-case analogical reasoning.

[36]   In full, the asserted claim read as follows:

A media system, comprising:

  a network based media managing system that maintains a library of content that a given user has a right to access and a customized user interface page for the given user;

  a collection of instructions stored in a non-transitory storage medium and configured for execution by a processor of a handheld wireless device, the collection of instructions operable when executed: (1) to initiate presentation of a graphical user interface for the network based media managing system; (2) to facilitate a user selection of content included in the library; and (3) to send a request for a streaming delivery of the content; and

  a network based delivery resource maintaining a list of network locations for at least a portion of the content, the network based delivery resource configured to respond to the request by retrieving the portion from an appropriate network location and streaming a representation of the portion to the handheld wireless device.

Affinity Labs, 838 F.3d at 1267–68.

confines the claim to a particular solution to an identified problem." *Id.* The claim merely

described "the concept of delivering user-selected media content to portable devices." *Id.*

In *IPA Technologies, Inc. v. Amazon.com, Inc.*, Amazon's second cited case, the District

of Delaware resolved an *Alice* challenge in Amazon's favor. 307 F. Supp. 3d 356 (D. Del. 2018).

The patents there involved "navigating an electronic data source by means of spoken language."

Id. at 359.[37] Although the plaintiffs in *IPA* asserted that their asserted claims provided specific

technological solutions and improvements, the court nevertheless found the claims to be

"directed to the abstract idea of transmitting electronic data to a user in response to a spoken

request from the user." *Id.* at 363. In support of this conclusion, the *IPA* court found that the

patents in question "fail[ed] to provide technological solutions to the problems they identif[ied]."

Id. at 364. This was so, the *IPA* court reasoned, because the patents identified "the complex

format of electronic databases as a problem for users" and identified "the goal of the invention,"

but then proceeded to claim "the objective of the invention itself." *Id.* This, in turn, was true

because the patent claims were "drafted so broadly as to cover any method that can achieve

---

[37]   The following claim serves as a representative example of the patents at issue in the *IPA* case:

> A method for speech-based navigation of an electronic data source, the electronic data source being located at one or more network servers located remotely from a user, comprising the steps of:
>
> (a) receiving a spoken request for desired information from the user;
>
> (b) rendering an interpretation of the spoken request;
>
> (c) constructing at least part of a navigation query based upon the interpretation;
>
> (d) soliciting additional input from the user, including user interaction in a non-spoken modality different than the original request without requiring the user to request said non-spoken modality;
>
> (e) refining the navigation query, based upon the additional input;
>
> (f) using the refined navigation query to select a portion of the electronic data source; and
>
> (g) transmitting the selected portion of the electronic data source from the network server to a client device of the user.

IPA, 307 F. Supp. 3d at 359.

navigating electronic databases by spoken natural language input." *Id.* All they recited, in

essence, were "the basic steps that would be required" to achieve that aim. *Id.*

 A cursory review of the claims at issue shows that the *IPA* court was right. *See supra*

n.37. Each step of the method claimed there proved conclusory and, indeed, self-evidently

necessary to achieve the invention's stated goal of allowing spoken language to serve as an input

for "rapidly searching and accessing desired content." *IPA*, 307 F. Supp. 3d at 364 (quoting the

specification). Any method of achieving that aim would necessarily include those steps:

> Any database search that begins with a request in a format not accepted by the
> database will require receipt, interpretation, and translation of the request to a
> format compatible with the database. Searching the database and transmitting the
> results to the user are required to retrieve any results from a database, even when
> searches are conducted entirely in a format compatible with the database.

Id. at 364 (citations omitted). Nothing included in any of the claims changed "the overall idea"

to which the claims were directed, and thus the *IPA* court found that the asserted patents failed

Alice Step One. *Id.* at 366.

 Amazon also cites to *USC IP Partnership v. Meta Platforms, Inc.* In that unpublished

case, the Federal Circuit analyzed the abstractness of a patent that claimed "a method for

determining the intent of a visitor to a webpage and using that intent to select and recommend

webpages to the visitor." *USC IP P'ship v. Meta Platforms, Inc.*, No. 2022-1397, 2023 WL

5606977, at *2 (Fed. Cir. Aug. 30, 2023).[38] The parties stipulated that "intent" meant "a unique

---

[38] Both the district and appellate court deemed the following claim representative:

A method for predicting an intent of a visitor to a webpage, the method comprising:

receiving into an intent engine at least one input parameter from a web browser displaying the
webpage;

processing the at least one input parameter in the intent engine to determine at least one
inferred intent;

providing the at least one inferred intent to the web browser to cause the at least one inferred
intent to be displayed on the webpage;

purpose or usage of the website." *Id.* The Federal Circuit agreed with the district court that the asserted claims were "directed to the abstract idea of 'collecting, analyzing and using intent data.'" *Id.* (quoting the district court). As a review of the claim language makes clear, *see supra* n.38, the claimed "intent engine" — the claim element that performed all of the work of determining intent — constituted nothing but a "purely functional 'black box.'" *Id.* The claims, the Federal Circuit concluded, failed to present "a technical solution to a technical problem," and nothing in the claims "affect[ed] the functionality of the computer itself." *Id.* at *3. The claims thus failed *Alice* Step One.

Amazon asserts that its next case, *Trinity Info Media, LLC v. Covalent, Inc.*, 72 F.4th 1355 (Fed. Cir. 2023), invalidated claims like those at issue here. (Amazon's Br. at 32 n.4.) There, the Federal Circuit concluded that the asserted claims were directed to "the abstract idea of matching based on questioning." 72 F.4th at 1361.[39] The *Trinity* court tasked itself with

---

> prompting the visitor to confirm the visitor's intent;
>
> receiving a confirmed intent into the intent engine;
>
> processing the confirmed intent in the intent engine to determine at least one recommended webpage that matches the confirmed intent, the at least one recommended webpage selected from a plurality of webpages within a defined namespace;
>
> causing the webpage in the web browser to display at least one link to the at least one recommended webpage;
>
> prompting the visitor to rank the webpage for the inferred intent;
>
> receiving a rank from the web browser; and
>
> storing a datapoint comprising an identity of the webpage, the inferred intent and the received rank.

USC IP, 2023 WL 5606977, at *2.

[39]   Each claim at issue in *Trinity Info* recited poll-based computer systems. The following claim is representative:

> A poll-based networking system, comprising:
>
> > a data processing system having one or more processors and a memory, the memory being specifically encoded with instructions such that when executed, the instructions cause the one or more processors to perform operations of:

"ascertain[ing] the basic character of the claimed subject matter," noting that the claims could

not be restated at such a high level that they "would be virtually guaranteed to be abstract." *Id.*

(cleaned up). A telling negative sign, the court said, was if "the claimed functions [were] mental

processes that can be performed in the human mind or using a pencil and paper." *Id.* at 1361–62

(cleaned up). Applying these principles, the *Trinity* court summarized the claims at issue as

requiring a basic series of steps: "(1) receiving user information; (2) providing a polling

question; (3) receiving and storing an answer; (4) comparing that answer to generate a

'likelihood of match' with other users; and (5) displaying certain user profiles based on that

likelihood." *Id.* at 1362. The focus of these steps, the court said, was "collecting information,

analyzing it, and displaying certain results," a classically patent-ineligible abstraction. *Id.*

(citation omitted). A human mind could easily perform these steps, so the court concluded that

the abstract idea of "matching based on questioning" had been claimed. *Id.*

   Dialect, of course, presents cases of its own. One case that found a computer-based claim

patent-eligible at *Alice* Step One is *DDR Holdings, LLC v. Hotels.com, L.P.*, 773 F.3d 1245 (Fed.

Cir. 2014). There, the Federal Circuit discussed patents that claimed "systems and methods of

generating a composite web page that combines certain visual elements of a 'host' website with

---

    receiving user information from a user to generate a unique user profile for the user;

    providing the user a first polling question, the first polling question having a finite set of answers and a unique identification;

    receiving and storing a selected answer for the first polling question; comparing the selected answer against the selected answers of other users, based on the unique identification, to generate a likelihood of match between the user and each of the other users; and

    displaying to the user the user profiles of other users that have a likelihood of match within a predetermined threshold.

Trinity Info, 72 F.4th at 1359.

content of a third-party merchant." *Id.* at 1248.[40] The *DDR* court found that these claims

survived the *Alice* gauntlet. It reasoned that the patent claimed a solution "necessarily rooted in

computer technology in order to overcome a problem specifically arising in the realm of

computer networks." *Id.* at 1257. The court emphasized that the claims at issue did not "broadly

and generically claim 'use of the Internet' to perform an abstract business practice" and that "the

claims at issue do not attempt to preempt every application of the idea of increasing sales by

making two web pages look the same." *Id.* at 1258–59. Thus, the claims proved eligible for

patent protection.

  In *Enfish, LLC v. Microsoft Corp.*, 822 F.3d 1327 (Fed. Cir. 2016), another of Dialect's

cases, the Federal Circuit found a different software claim non-abstract at *Alice* Step One. The

Enfish court analyzed claims on the manipulation of computer memory using self-referential

logical tables.[41] The district court in *Enfish* characterized the claims as directed to the abstract

---

[40] The following claim was cited as representative:

 An e-commerce outsourcing system comprising:

 a) a data store including a look and feel description associated with a host web page having a
   link correlated with a commerce object; and

 b) a computer processor coupled to the data store and in communication through the Internet
   with the host web page and programmed, upon receiving an indication that the link has been
   activated by a visitor computer in Internet communication with the host web page, to serve a
   composite web page to the visitor computer wit[h] a look and feel based on the look and feel
   description in the data store and with content based on the commerce object associated wit[h]
   the link.

DDR, 773 F.3d at 1249.

[41] The following claim illustrates what the *Enfish* court had in mind:

 A data storage and retrieval system for a computer memory, comprising:

  means for configuring said memory according to a logical table, said logical table including:

   a plurality of logical rows, each said logical row including an object identification
   number (OID) to identify each said logical row, each said logical row corresponding to a
   record of information;

concept of "organizing information using tabular formats." *Id.* at 1336–37. The Federal Circuit disagreed and found that the claims were "not simply directed to *any* form of storing tabular data" but rather "are specifically directed to a *self-referential* table for a computer database." *Id.* at 1337. It constituted "a specific type of data structure designed to improve the way a computer stores and retrieves data in memory." *Id.* at 1339. It was thus entirely unlike cases "where general-purpose computer components are added post-hoc" to an abstract idea to enable the patentee to loophole their way around the *Alice* doctrine. *Id.* The claims, in short, were not abstract.

In the last of Dialect's cited cases that the Court finds helpful, *Visual Memory LLC v. NVIDIA Corp.*, 867 F.3d 1253 (Fed. Cir. 2017), the Federal Circuit found yet another software claim non-abstract. There, the Federal Circuit held, the claims at issue were "directed to an improved computer memory system, not to the abstract idea of categorical data storage." *Id.* at 1259.[42] The claims did not "recite all types and all forms of categorical data storage," preempting the abstract idea entirely. *Id.* Instead, the patent's specification explained that

---

> a plurality of logical columns intersecting said plurality of logical rows to define a plurality of logical cells, each said logical column including an OID to identify each said logical column; and
>
> means for indexing data stored in said table.

Enfish, 822 F.3d at 1336.

[42]   The following claim from the case proves illustrative:

> A computer memory system connectable to a processor and having one or more programmable operational characteristics, said characteristics being defined through configuration by said computer based on the type of said processor, wherein said system is connectable to said processor by a bus, said system comprising:
>
> > a main memory connected to said bus; and
> >
> > a cache connected to said bus;
> >
> > wherein a programmable operational characteristic of said system determines a type of data stored by said cache.

Visual Memory, 867 F.3d at 1257.

"multiple benefits flow[ed]" from the patent's system, suggesting to the *Visual Memory* court

that the claims were "directed to a technological improvement" instead of reciting generic steps

using conventional computing components. *Id.* at 1259–60. The patent thus survived at Step

One.

A.   Application of the Relevant Law

The preceding section's exposition, though protracted, offers a robust set of comparable

cases that can now power the analogical reasoning that precedent dictates. With the Court's

toolkit now fully arrayed, the Court can and does make short work of *Alice* Step One.

1.   The '006 Patent

Claim 1 of the '006 Patent does not target an abstract idea, and neither does Claim 5.[43]

The methods disclosed in those claims prove to be much more analogous to the particularized

and computer-oriented ones approved of in *DDR*, *Enfish* and *Visual Memory* than the abstract

and conclusory ones invalidated in *IPA*, *USC IP* or *Trinity Info*.

Claims 1 and 5 do not claim "the abstract idea of understanding and responding [to] a

spoken request," nor do they purport to do so. (Amazon's Br. at 30.) The claims, in their

preambles, disclose methods "method for processing natural language speech utterances." '006

Pat. col. 25 ll. 49–50 (Claim 1).[44] The method that they set out does not come close to claiming

the entire universes of understanding and responding to speech. Instead, they claim only a

particular method. *See supra* § II.A (summarizing that method). To be sure, the software that

performs that method is somewhat abstract; Amazon correctly accuses the '006 Patent of lacking

---

[43]   Because the Court will grant Amazon summary judgment of noninfringement on Claim 10 of the '006 Patent, the Court will not address how the abstract idea exception applies to that claim. *See supra* § IV.B.3.

[44]   The Court treats Claim 1 of the '006 Patent as representative. All of the Court's statements concerning Claim 1 apply *a fortiori* to Claim 5.

any indication of what a computer programmer ought to put inside of, say, a "domain agent" to make it "process[] the formulated request." '006 Pat. col. 26, l. 17–19. But understanding speech using context does not necessarily require deploying dynamically updated "dictionary and phrase tables," identifying the person speaking, inferring and extracting parameters, or even breaking down an utterance into "recognized words or phrases." *See generally id.* Unlike the patents found invalid in *IPA* and *USC IP*, Claims 1 and 5 do not capture "any method" that can achieve its stated goal, *IPA*, 307 F. Supp. 3d at 364, nor do they fail to "affect the functionality of the computer itself," *USC IP*, 2023 WL 5606977, at *3.

Neither do Claims 1 and 5 recite processes "that can be carried out by a human." (Amazon's Br. at 31.) True, the claimed results — "recognizing speech, using context to understand the speaker's intent, and formulating a response" — are "what every human does in a conversation." (*Id.*) But Amazon confuses the claim's goal or result with the method that the claim uses to achieve it. As the case law discussed in the previous section demonstrates, a patent's claimed method must itself constitute the abstract goal or result claimed for the patent to be directed to an abstract idea. Amazon cannot and does not argue that the tables, agents, parsers and so forth that populate the claims simply restate how human cognition works. If that were so, the inventors of the '006 Patent would have achieved an astonishing advance in cognitive science. It should come as no surprise that Claims 1 and 5 set their sights on a fundamentally human activity; programming computers to imitate human behaviors constitutes the precise problem to which the '006 Patent attempts to provide a solution. In addition, Amazon's argument that the claimed computer "act[s] like a human" because it "understand[s] and respond[s] using 'context'" has no purchase following the Court's *Markman* Opinion. (*Id.*; Amazon's Reply at 19 & n.12.) There, the Court explained at length and in detail, at Amazon's

57

urging, that the Asserted Patents do not give the word "context" its "plain and ordinary meaning." *Markman Op.*, 2024 WL 1859806, at *6–8. Amazon cannot suddenly double back and assert that "context" carries its ordinary meaning to obfuscate a meaning that its own *Markman* briefs helped to make clear.

To summarize, Claims 1 and 5 of the '006 Patent disclose specific processes for improving the functionality of a computer, rather than conclusory and computer-divorced steps that invoke technology merely as a tool. They thus claim patentable subject matter.[45]

2.   The '720 Patent

Much of the Court's analysis of the '006 Patent applies to the '720 Patent as well. Amazon asserts that the '720 Patent claims "the same abstract idea" as the '006 Patent — understanding and responding to a spoken request. (Amazon's Br. at 35.) However, as Dialect correctly insists, Amazon describes the '720 Patent at far too high a level of abstraction. (Dialect's Opp. at 34 (citing and quoting *McRO*, 837 F.3d at 1313).) Just as with the '006 Patent, the '720 Patent does far more than parrot back the steps needed to understand and respond to spoken words. Like the patent considered in *Enfish*, the '720 Patent does not claim any and all methods by which a spoken request might be understood and responded to, and it likewise does not use generic computer components to dress up what ought to be an analog process. 822 F.3d at 1337, 1339. The '720 Patent's claims stand deeply "rooted in computer technology" to overcome a computer-specific problem. *DDR*, 773 F.3d at 1257. They cannot be compared to

---

[45]   Because Amazon's *Alice* arguments regarding the dependent claims of the '006 Patent depend on the Court first finding that Claims 1 and 5 disclose an abstract idea, those arguments can be disregarded without separate analysis. (*See* Amazon's Br. at 31–32.)

the sort of "just-do-it-on-a-computer" claims invalidated in Amazon's cited cases. *See supra*

nn.36–39. Dialect's asserted claims thus do not run afoul of § 101's abstract idea exception.[46]

3.   The '957 Patent

Amazon charges the '957 Patent with the same sort and degree of abstraction that it

ascribes to the '006 and '720 Patents. For similar reasons to those stated above, the Court

disagrees that the '957 Patent claims patent-ineligible subject matter. A brief discussion of

Claim 1, the patent's only substantively independent claim, *see supra* n.12, suffices to explain

why.

Like the '006 and '720 Patents, the '957 Patent aims to enable computer technology that

simplifies interactions with humans. Because computers cannot process speech like people do,

the '957 Patent "uses context, prior information, domain knowledge, and user specific profile

data to achieve a natural environment for users" to communicate with machines. '957 Pat. col. 2,

l. 65–col. 3, l. 1. Claim 1, for its part, describes a system that understands and responds to a

spoken request by using a particular process: breaking the request down into individual words,

comparing those words to bits of data to identify the subject matter of the request, and then using

that subject matter data to assign meaning to the spoken request. *See supra* § II.C. Even when

described at this high level of generality, Claim 1 does not simply "describe a desired function or

outcome." *Affinity Labs*, 823 F.3d at 611. Nor can the steps of the method that Claim 1 recites

be performed intuitively "in the human mind." *Trinity Info*, 72 F.4th at 1361–62. Described at

the appropriate level of generality, Claim 1 — and thus the '957 Patent, *see supra* nn.45, 46 —

appears distinctly non-abstract at *Alice* Step One.

---

[46]   Just as with the '006 Patent, the dependent claims of the '720 Patent survive, because
they are directed to the same non-abstract method and system as the independent claims that they
incorporate. *See supra* n.45.

4.   The '468 Patent

Next, the Court addresses the '468 Patent, which Amazon asserts is directed to the abstract idea of understanding and responding to a multimodal request. (Amazon's Br. at 37–38.) For similar reasons to those discussed above, the Court finds that Amazon describes the claims at far too high a level of generality. Grounding its analysis in the text of the relevant claims, the Court finds that the '468 Patent does not monopolize the very idea of processing a multimodal input.

The Court centers its analysis on Claim 19, the only independent claim of the '468 Patent that Dialect asserts. *See supra* § II.D. A review of the Court's summary of Claim 19, *id.*, and representative claims found abstract at *Alice* Step One, *see supra* § V.A.3, makes clear that Claim 19 has more specificity and less preclusive effect than the patents at issue in *Affinity Labs*, *IPA* or *Trinity Info*. Amazon asserts that Claim 19 is directed to a mere result, because it describes no more than "what a human does in everyday multimodal conversations: merging the information received and relying on context." (Amazon's Br. at 38.) However, the Court must express some skepticism that the many steps of Claim 19 — transcribing a non-speech input, applying three distinct knowledge-based models to interpret a speech input, merging the two, identifying a context, referring the combined input to a domain agent, and then outputting the result — recite simply "what a human does in everyday multimodal conversation." Humans, admittedly, use context, in the ordinary sense of that word, to interpret language. But the Court has already determined that the word "context" in Claim 19 does not carry its plain and ordinary meaning, and in any event, Claim 19 does far more than simply determine context and apply it.

Here, as with the '006, '720 and '957 Patents, Amazon overgeneralizes the claims to erase precisely the features that render them non-abstract — their recitation of a particular method designed for the computer environment that aims for the goal of interpreting language

without claiming "the objective of the invention itself." *IPA*, 307 F. Supp. 3d at 364. The Court

finds that the '468 Patent passes the *Alice* test at Step One.

        5.   **The '039 Patent**

The Court finally turns to the last patent asserted, the '039 Patent. Amazon argues that

this patent claims the same abstract idea as the '468 Patent, namely, the idea of understanding

and responding to a multimodal request. (Amazon's Br. at 37–38.) The Court begins by

analyzing Claim 13, the only independent claim of the '039 Patent that Dialect asserts. The

claim appears vague, and it lies the closest to abstraction of the many claims and patents asserted

in this case. Close analysis, however, reveals that Claim 13 is not directed to an abstract idea.

Claim 13 recites a cursory and somewhat conclusory set of steps for "processing speech

and non-speech communications." '039 Pat. col. 30, ll. 39–40; *see supra* § II.E. Simple as those

steps may be, the Court finds that Claim 13 offers more than Amazon's proffered abstract idea of

"understanding and responding to a multimodal request" to overgeneralize the terms of the

claim. (Amazon's Br. at 39.) Amazon asserts that Claim 13 recites no technology for using

context to respond to a multimodal request and thus discloses no method "to achieve this

aspirational result in a non-abstract way." (Amazon's Br. at 38 (citing *Two-Way Media Ltd. v.*

Comcast Cable Commc'ns, LLC, 874 F.3d 1329, 1337 (Fed. Cir. 2017).) For the following

reasons, the Court disagrees.

First, Amazon's invocation of *Two-Way Media* fails. The claims that the Federal Circuit

found invalid in *Two-Way Media* "require[d] [] functional results" without providing any

concrete means for achieving them. 874 F.3d at 1337–38. A brief analysis of just one claim

reveals why this was so: The claimed method's steps recited aspirational results without

requiring any particular means, technique, or tool to be used, resulting in a claim directed

towards a bare result. *Id.* at 1334–35, 1338.[47] Admittedly, several steps of the '039 Patent

constitute such result-oriented limitations. For example, the first two elements recite "receiving

the speech and non-speech communications" and "transcribing" them to create two messages.

'039 Pat. col. 30, ll. 41–44. Those elements describe aspirational goals and would be met by any

method, no matter how simple or complex, that accomplishes the stated goal. However, the

process that follows — searching for text combinations, comparing them to entries, computing

relevance scores based on those comparisons, and referring the message for analysis — describes

means, not ends. These limitations, labeled as elements (d) to (h) in the Court's initial

description of the claim, *see supra* § II.E, constitute the focus of the claimed invention. However

those claim elements may be phrased, they do not constitute a necessary part of the process of

understanding multimodal inputs, nor do they restate what humans do when they seek to

understand communications by others.[48] The focus of Claim 13 thus proves non-abstract in

nature, so the '039 Patent survives at *Alice* Step One.

---

[47]   For purposes of comparison, a representative claim at issue in *Two-Way Media* read as
follows:

> A method for transmitting message packets over a communications network comprising the steps
> of:
>
> > converting a plurality of streams of audio and/or visual information into a plurality of streams
> > of addressed digital packets complying with the specifications of a network communication
> > protocol,
> >
> > for each stream, routing such stream to one or more users,
> >
> > controlling the routing of the stream of packets in response to selection signals received from
> > the users, and
> >
> > monitoring the reception of packets by the users and accumulating records that indicate which
> > streams of packets were received by which users, wherein at least one stream of packets
> > comprises an audio and/or visual selection and the records that are accumulated indicate the
> > time that a user starts receiving the audio and/or visual selection and the time that the user
> > stops receiving the audio and/or visual selection.

Two-Way Media, 874 F.3d at 1334–35.

[48]   For this reason, Amazon's citation of *Trinity Info* and *Affinity Labs* has no force.
(Amazon's Br. at 38.) True, result-oriented claims cannot be patented. (*Id.* (citing *Trinity Info*,

Moreover, Judge Ellis's *Alice* Opinion does not require the opposite result, despite

Amazon's insistence that his holding "applies with equal force here." (Amazon's Br. at 38.)

There, Judge Ellis analyzed U.S. Patent No. 9,031,845 and, after applying the *Alice* test, found it

impermissibly abstract. *Alice Op.*, 2023 WL 7381551, at *1 & n.1. Judge Ellis, considering the

entirety of the relevant claim text, found it to be "plainly result-oriented," as it was "directed to

any arrangement of programs and processors" that accomplished its "stated goal." *Id.* at *4.[49]

The same reasoning does not apply here. To begin with, the '845 Patent expressly claimed "one

or more physical processors . . . programmed to execute one or more computer program

instructions which, when executed cause the one or more physical processors to" perform a

series of steps. The claim thus targeted no particular arrangement of components and no

---

72 F.4th at 1362 and *Affinity Labs*, 838 F.3d at 1269).) But this claim, when read as a whole, is
not oriented towards a result.

[49]   The relevant claim text follows:

A mobile system for processing natural language utterances, comprising:

one or more physical processors at a vehicle that are programmed to execute one or more
computer program instructions which, when executed, cause the one or more physical
processors to:

receive a natural language utterance associated with a user;

perform speech recognition on the natural language utterance;

parse and interpret the speech recognized natural language utterance;

determine a domain and a context that are associated with the parsed and interpreted
natural language utterances;

formulate a command or query based on the domain and the context;

determine whether the command or query is to be executed on-board or off-board the
vehicle;

execute the command or query at the vehicle in response to a determination that the
command or query is to be executed on-board the vehicle; and

invoke a device that communicates wirelessly over a wide area network to process the
command or query such that the command or query is executed off-board the vehicle in
response to a determination that the command or query is to be executed off-board the
vehicle.

'845 Pat. col. 32, ll. 30–57.

particular set of techniques. So long as the end was achieved, the '845 Patent covered it. And

additionally, the relevant claim text provided nothing in the way of a *means* for achieving its goal

of processing natural language other than a bare command to "determine a domain and a

context." Each step of the claimed method provided nothing to distinguish the claimed

"method" from the overall goal. As the Court has explained, those flaws cannot be found in the

'039 Patent.

Judge Ellis correctly determined that Claim 1 of the '845 Patent was directed towards an

abstract idea. Judge Ellis, just as correctly, chose not to decide the remainder of the *Alice*

question, recognizing that the remaining Asserted Patent presented closer questions. Applying

the same reasoning and the same test that Judge Ellis used, the Court now concludes that the

'039 Patent does not lay claim to an abstract idea.[50]

---

[50]   As before, because the sole independent claim of the '039 Patent that Dialect asserts
survives the *Alice* analysis, the Court need not individually interrogate each of the dependent
claims that follow from it. *See supra* nn.45, 46.

VI. CONCLUSION

The consequences of the Court's analysis follow.

First, Count III of Dialect's Amended Complaint must be dismissed. That count pleads infringement of the '327 Patent, but Alexa does not infringe that patent as a matter of law.
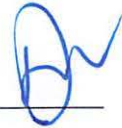
Second, Alexa does not infringe Claims 10 or 11 of the '006 Patent as a matter of law, but infringement of Claims 1, 2, 3 and 5 presents issues of fact that must be tried to a jury.

Third, Dialect presents genuine questions of fact that preclude summary judgment of noninfringement on any asserted claim of the '720, '957, '468 and '039 Patents.

Fourth and last, the asserted claims of the '006, '720, '957, '468 and '039 Patents claim patent-eligible subject matter rather than abstract ideas.

An appropriate Order shall issue. Let the Clerk file a copy of this Memorandum Opinion and notify all counsel of record.

/s/

David J. Novak
United States District Judge

Alexandria, Virginia
Date: July 30, 2024

65